NNT : 2015SACLS217

Thèse de Doctorat
de
l'Université Paris-Saclay
préparée à
l'Université Paris-Sud

École Doctorale N° 574
École doctorale de mathématiques Hadamard

Spécialité de doctorat : Mathématiques fondamentales

Par

Mme Daphné Dieuleveut

Coupe et reconstruction d'arbres et de cartes aléatoires

Thèse présentée et soutenue à Orsay, le 10 décembre 2015

Après avis des rapporteurs :

M. Jean-François Marckert	Directeur de recherche	Université Bordeaux 1
M. Anton Wakolbinger	Professeur	Goethe-Universität Frankfurt

Composition du Jury :

Mme Bénédicte Haas	Professeur	Université Paris-Saclay	Présidente
M. Jean-François Marckert	Directeur de recherche	Université Bordeaux 1	Rapporteur
M. Nicolas Broutin	Professeur	Inria	Examinateur
M. Nicolas Curien	Professeur	Université Paris-Saclay	Examinateur
M. Yves Le Jan	Professeur	Université Paris-Saclay	Directeur de Thèse
M. Grégory Miermont	Professeur	ENS de Lyon	Directeur de Thèse

Table des matières

Remerciements

En premier lieu, je tiens à exprimer mon immense reconnaissance envers mes directeurs : Grégory Miermont, pour l'esthétique des sujets qu'il m'a proposés au début et au cours de cette thèse, ainsi que sa constante implication, sa disponibilité, et pour m'avoir considérablement aidée à améliorer la rigueur et la clarté de ce manuscrit ; et Yves Le Jan, pour son suivi plus ponctuel mais tout aussi enrichissant, et ses conseils avisés.

Je remercie Jean-François Marckert et Anton Wakolbinger d'avoir accepté de rapporter cette thèse, et de s'être acquittés de cette tâche avec un très grand soin. Merci également à Nicolas Broutin, Nicolas Curien et Bénédicte Haas, d'avoir accepté avec autant d'enthousiasme de faire partie du jury. C'est un honneur pour moi de pouvoir leur présenter mes travaux en cette occasion.

Au cours de ma troisième année, j'ai eu la chance de pouvoir échanger de manière régulière avec Louigi Addario-Berry et Christina Goldschmidt, en particulier lors d'un bref séjour à Oxford. Merci à eux de m'avoir accueillie avec autant d'amabilité et d'ouverture, et pour les discussions que nous avons partagées, sans lesquelles le chapitre 3 de cette thèse n'aurait pas vu le jour.

Les "journées cartes" et l'école d'été de Saint-Flour ont aussi été pour moi d'importantes occasions d'échanger avec des personnes travaillant sur des sujets proches des miens. Je remercie leurs organisateurs et participants d'avoir non seulement rendu possibles ces échanges, mais surtout fait de leur mieux pour qu'ils se déroulent dans des conditions optimales.

Enfin, l'excellence des professeurs dont j'ai suivi l'enseignement en probabilités a eu une grande influence dans le choix de ce domaine pour ma thèse. Je tiens en conséquence à leur adresser des remerciements particuliers, notamment à Jean Bertoin, Wendelin Werner et Jean-François Le Gall [1].

Avant de passer à des louanges moins académiques et plus amicales, je remercie également les personnels administratifs de l'université d'Orsay (puis de Paris-Saclay), et de l'ENS de Lyon, qui ont su faire en sorte que mes déplacements réguliers se déroulent dans les meilleures conditions possibles. Merci en particulier à Valérie Blandin-Lavigne, Catherine Ardin, Magalie Le Borgne, Virginia Goncalves, et Sandy Artero.

Je ne saurais aborder cette seconde salve de remerciements sans une pensée émue pour un *grand* nombre des professeurs qui ont accompagné mon parcours, bien avant mon arrivée à l'ENS. Pour ne citer que les plus essentiels, je me rappellerai toujours la bienveillance de Mme Laguionie, la première à m'avoir parlé de "maths sup' ", le regard de M. Hébrard lorsqu'il abordait un nouveau chapitre ou dessinait de jolies intégrales au tableau, la rigueur de l'enseignement de M. Blanchard et la richesse de celui de M. Canonico [2]... ce qui ne signifie pas que j'oublie les professeurs d'autres matières, autant littéraires que scientifiques, qui ont su nourrir l'envie d'apprendre qui m'a menée ici.

Durant ces trois années de thèse, j'ai évolué dans l'environnement convivial des laboratoires de mathématiques d'Orsay et de l'ENS de Lyon. Je remercie les doctorants de ces deux équipes — les premiers, pour avoir toléré mes absences régulières sans trop m'oublier, et les seconds, pour m'avoir très bien accueillie lors de mes passages bi-mensuels. Merci en particulier à Élodie, Céline, Émilie et Valérie R. pour leurs sourires ; à Loïc, Sébastien et Romain pour nos discussions plus ou moins *cruchiales* au self de Lyon ; et à Vincent B., pour quelques TD d'algèbre et beaucoup de clous.

J'adresse également une pensée à tous les jeunes probabilistes avec lesquels j'ai eu le plaisir de partager des discussions mathématiques, et notamment à Erich, Robin S., Igor, Jérémie et Cyril.

En matière de discussions mathématiques, le séminaire de l'abbé Mole fut également un lieu d'échanges particulièrement stimulant, sans doute en raison des étroits liens d'amitié qui

1. et à ses polys.
2. et le caractère indispensable de son Canogoshi.

s'ajoutent à nos passions communes. Je remercie donc tous les membres du C6, en me permettant de mélanger allègrement citoyens de la république montrougienne et participants du séminaire sus-cité : Arnaud, Aurélie, Bastien, Benoît, David, Delphine, Fathi, Gabriel, Irène, Jack, Julia, Julie, Julien, Manon, Marie A., Marie D., Martin, Nicolas L., Pierre, Vincent P. et Yasmine. (En espérant que celles qui ne font partie de l'un des deux groupes que par alliance excuseront la confusion, et que ceux que j'ai déjà prévu de citer ailleurs me pardonneront l'omission.)

Merci aux membres du B.C. qui m'ont régulièrement *soutenu les coudes* : Amélie, Benjamin, Christine, Denis, Iryna, Juliette, Kim, Nicolas H., Renaud, Robin P. (au cas où il réapparaîtrait un jour ?) et Vincent V.

Merci à Quentin pour nos cathartiques séances de dessin et de chant, et à tous les autres bridgeurs de l'ENS (notamment Olivier, Baptiste et Adeline) pour avoir supporté ces digressions bruyantes avec le sourire.

Merci à Catherine, Silvain, Max et Camille pour de régulières aventures gourmandes et *hobbittesques*, et pour tous les excellents moments passés ensemble.

Merci à la famille de Rémi de toujours m'accueillir à bras ouverts, et notamment à Laurence, pour avoir partagé son expérience en matière de découpage d'arbres à la tronçonneuse, et à François, pour m'avoir initiée à la subtilité des cartes (plus ou moins aléatoires) de Rail.

Last but not least, merci à toute ma famille pour leur soutien indispensable. À Jordane, qui supporte avec le sourire nos bêtises. À Yvonne, avec des vœux de prompt rétablissement. À Roland, qui suit de loin mais toujours avec un regard positif. À Aymeric, qui répond avec une extraordinaire ponctualité à mes messages dès qu'il y a des maths dedans. À Anouk, Floriane et Dominique, qui à défaut de relire mes hiéroglyphes, font vivre par leurs relectures un autre monde qui me tient tout autant à cœur. À Rémi, enfin, sans qui la vie serait restée grise.

Chapitre 1

Introduction

1 Fragmentations d'arbres aléatoires et arbres des coupes associés

La première partie de cette thèse est consacrée à l'étude de l'*arbre des coupes* associé à différentes fragmentations d'arbres aléatoires discrets et continus.

1.1 Arbres aléatoires discrets et continus

Les arbres de Galton-Watson sont l'un des modèles d'arbres aléatoires discrets les plus connus ; ils permettent notamment de décrire de manière naturelle la généalogie d'une population où les individus se reproduisent indépendamment et suivant la même loi. On appelle *limites d'échelle* de tels arbres les arbres continus obtenus lorsque l'on fait tendre la taille de la population (c'est-à-dire le nombre de nœuds par lequel on conditionne) vers l'infini, en multipliant la longueur des arêtes par un facteur adapté qui tend vers zéro. Dans le cas d'une loi de reproduction de variance finie, la limite d'échelle est l'arbre brownien d'Aldous, étudié en détail dans [5, 6, 7]. Dusquesne et Le Gall [32, 33] ont introduit une classe plus générale d'arbres aléatoires continus, les arbres stables, qui correspondent également à des limites d'échelle d'arbres de Galton-Watson, si la loi de reproduction est dans le domaine d'attraction d'une loi stable. Nos travaux se concentrent sur ces deux modèles.

Nous commençons par fixer un cadre commun pour étudier les arbres discrets et continus, afin d'expliciter les notions de convergence utilisées ; nous donnons ensuite les définitions des arbres de Galton-Watson, de l'arbre brownien et des arbres stables.

1.1.1 Arbres réels et notion de convergence

Un arbre discret est défini comme un graphe connexe, sans cycle. La notion d'arbre réel étend cette définition à un cadre continu, en conservant la propriété essentielle : l'existence d'un unique chemin injectif reliant tout couple de points. Formellement, un arbre réel (ou $\mathbb{R}$-arbre) est un espace métrique (T, d) tel que, pour tous points $u, v \in T$, les deux conditions suivantes sont vérifiées :

— Il existe une isométrie $\varphi_{u,v}$ de $[0, d(u, v)]$ dans T telle que $\varphi_{u,v}(0) = u$ et $\varphi_{u,v}(d(u, v)) = v$.

— Pour tout chemin injectif φ de $[0, 1]$ dans T tel que $\varphi(0) = u$ et $\varphi(1) = v$, on a

$$\varphi([0, 1]) = \varphi_{u,v}([0, d(u, v)]) =: [\![u, v]\!].$$

On considère généralement des arbres enracinés, c'est-à-dire munis d'un point distingué qu'on appelle la racine.

Les arbres réels que nous étudierons peuvent être définis *via* un codage par une fonction continue, en utilisant une construction due à Aldous [7] (voir également [33, Section 2]). Nous allons donner un aperçu de cette construction, en partant du cas discret.

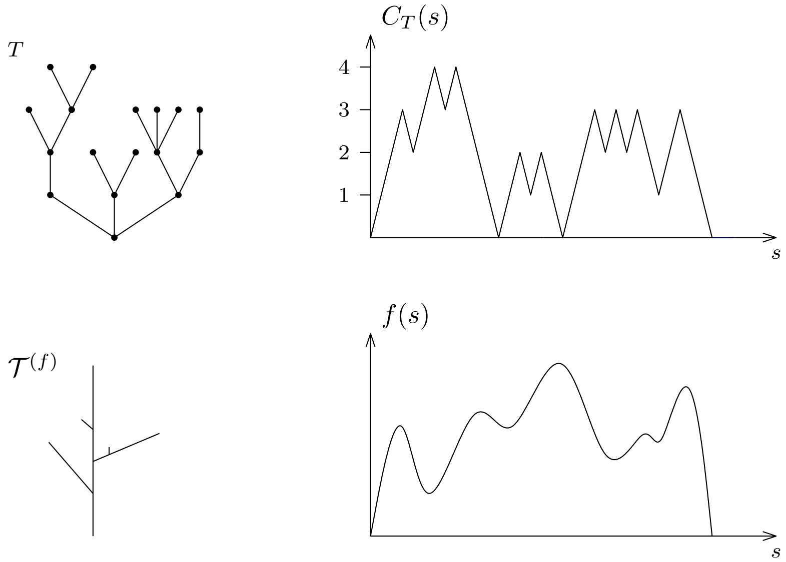

FIGURE 1.1 – Deux exemples illustrant la correspondance entre un arbre et sa fonction de contour : en haut, dans le cas d'un arbre discret, et en bas, pour un arbre codé par une fonction continue.

Soit T un arbre discret enraciné, c'est-à-dire muni d'un sommet distingué, la racine ; on notera par la suite $\mathbf{T}$ l'ensemble de ces arbres. On fixe un plongement de T dans le plan, tel que chaque arête soit de longueur 1. À l'arbre T est alors associée une *fonction de contour* C_T, définie informellement de la manière suivante. On considère une particule qui fait le tour de l'arbre en partant de la racine, de gauche à droite, à vitesse 1. Pour tout $s \in [0, 2\#E(T)]$, on note $C_T(s)$ la hauteur de la particule à l'instant s (c'est-à-dire la distance la séparant de la racine). Il est intuitivement clair que l'on peut retrouver l'arbre T à partir de C_T. La construction d'Aldous définit de manière plus générale l'arbre codé par une fonction donnée. Soit f une fonction continue, positive, définie sur un intervalle $[0, \varsigma]$ et nulle en 0 et ς. On considère la pseudo-distance $d^{(f)}$ définie par

$$d^{(f)}(s, s') = f(s) + f(s') - 2 \min_{[s \wedge s', s \vee s']} f \qquad \forall s, s' \in [0, \varsigma].$$

L'arbre réel $\mathcal{T}^{(f)}$ associé à f est l'espace obtenu en quotientant l'intervalle $[0, \varsigma]$ par la relation d'équivalence $d^{(f)}(\cdot, \cdot) = 0$. La figure 1.1 donne un exemple de cette construction. L'arbre $\mathcal{T}^{(f)}$ est enraciné au point $p^{(f)}(0)$, avec $p^{(f)}$ la projection canonique de $[0, \varsigma]$ sur $\mathcal{T}^{(f)}$. En outre, il est naturellement muni d'une mesure de probabilité $\mu^{(f)}$ définie comme la mesure image de la mesure uniforme sur $[0, \varsigma]$ par $p^{(f)}$. On appelle $\mu^{(f)}$ la *mesure de masse* sur l'arbre $\mathcal{T}^{(f)}$.

Notons qu'un élément de $\mathbf{T}$ peut également être vu comme un arbre réel mesuré, en le munissant de la mesure uniforme sur ses sommets. La topologie avec laquelle nous allons travailler est adaptée à cette structure d'arbre mesuré et enraciné. Le cadre général est le suivant.

On appelle espace métrique mesuré pointé tout quadruplet (X, d, μ, ρ), où μ est une mesure de probabilité borélienne sur l'espace métrique (X, d), et ρ est un point de X. De tels espaces (X, d, μ, ρ) et (X', d', μ', ρ') sont dits équivalents s'il existe une application mesurable ϕ de X dans X', telle que :

— $\phi(\rho) = \rho'$.

— pour tous $x, y \in \operatorname{supp}(\mu) \cup \{\rho\}$, on a $d(x, y) = d'(\phi(x), \phi(y))$,

— μ' est la mesure image de μ par ϕ,

On note $\mathbb{M}$ l'ensemble des classes d'équivalences d'espaces métriques mesurés pointés (X, d, μ, ρ) tels que (X, d) est complet, et $X = \operatorname{supp}(\mu) \cup \{\rho\}$. On appelle arbre continu tout élément de $\mathbb{M}$ correspondant à la classe d'équivalence d'un arbre réel mesuré *compact*. On munit l'ensemble $\mathbb{M}$

de la topologie de Gromov-Prokhorov pointée, pour laquelle la notion de convergence est définie de la manière suivante. Soient $\mathbf{X}_n = (X_n, d_n, \mu_n, \rho_n)$ des espaces métriques mesurés pointés, pour $n \in \mathbb{N} \cup \{\infty\}$. Pour tout n, on note $\xi_n(0) = \rho_n$, et $(\xi_n(i))_{i\geq 1}$ une suite d'éléments i.i.d. de X_n, de loi μ_n. On dit que la suite $(\mathbf{X}_n)$ converge vers $\mathbf{X}_\infty$ au sens de Gromov-Prokhorov si pour tout $k \in \mathbb{N}$, on a la convergence en loi

$$(d_n(\xi_n(i), \xi_n(j)))_{0\leq i,j\leq k} \xrightarrow[n\to\infty]{(\mathcal{L})} (d_\infty(\xi_\infty(i), \xi_\infty(j)))_{0\leq i,j\leq k} .$$

L'ensemble $\mathbb{M}$, muni de la topologie de Gromov-Prokhorov, est un espace polonais. Notons que lorsque les espaces $\mathbf{X}_n$, $n \in \mathbb{N} \cup \{\infty\}$ que l'on considère sont compacts, on peut également envisager une notion plus forte de convergence, appelée convergence au sens de Gromov-Hausdorff-Prokhorov. On renvoie au livre de Gromov [37, Chapitre 3] et à l'article [36] pour une étude détaillée de ces topologies.

Notons une propriété importante pour les arbres réels codés par une fonction de contour. Pour toute suite de fonctions f_n continues, positives, définies sur un même intervalle $[0, \varsigma]$, nulles en 0 et ς, la convergence

$$\|f_n - f\|_\infty \xrightarrow[n\to\infty]{} 0$$

implique la convergence de $\mathcal{T}^{(f_n)}$ vers $\mathcal{T}^{(f)}$ au sens de Gromov-Prokhorov (ainsi qu'au sens de Gromov-Hausdorff-Prokhorov). On aura généralement ce type de convergence à un facteur d'échelle près ; on note donc $a \cdot (T, d, \mu, \rho) = (T, a \cdot d, \mu, \rho)$ pour tout réel positif a.

1.1.2 Les arbres de Galton-Watson et leurs limites d'échelle

Les principaux modèles d'arbres aléatoires avec lesquels nous travaillerons sont les arbres de Galton-Watson et leurs limites d'échelles, l'arbre continu brownien et l'arbre stable.

Fixons une loi de probabilité ν sur $\mathbb{N}$. Un arbre de Galton-Watson de loi de reproduction ν correspond à l'arbre généalogique d'une population dans laquelle chaque individu se reproduit de manière indépendante, selon la loi ν. Cela se traduit par exemple par la caractérisation suivante :

— Le nombre d'enfants N de la racine suit la loi ν.

— Pour tout $n \geq 0$, sachant $N = n$, les n arbres formés par les enfants de la racine et leurs descendants sont des arbres de Galton-Watson indépendants, de loi de reproduction ν.

On travaillera en général avec une loi de reproduction critique (i.e. d'espérance égale à 1), telle que $\nu(\{1\}) \neq 1$, de manière à assurer que les arbres aléatoires correspondants aient presque sûrement un nombre fini de sommets.

Pour tout n, on note T_n un arbre de Galton-Watson de loi de reproduction ν, conditionné à avoir exactement n arêtes (on travaille implicitement pour les valeurs de n telles que cet événement a une probabilité non nulle). Comme indiqué précédemment, T_n est vu comme un arbre réel, muni de la distance de graphe et de la mesure uniforme sur ses sommets. Les théorèmes ci-dessous donnent la limite d'échelle de T_n dans les deux cas suivants :

1. La loi ν est de variance finie σ^2.
2. La loi ν appartient au bassin d'attraction de la loi stable d'indice α, avec $\alpha \in (1, 2]$.

(Notons que le premier cas peut être vu comme un cas particulier du second, pour $\alpha = 2$.)

Théorème (Aldous [7]). *Dans le cas 1, on a la convergence*

$$\left(\frac{\sigma}{\sqrt{n}} C_{T_n}(2nt)\right)_{0\leq t\leq 1} \xrightarrow[n\to\infty]{(\mathcal{L})} (2\mathbf{e}_t)_{0\leq t\leq 1} ,$$

avec $(\mathbf{e}_t)_{0\leq t\leq 1}$ *l'excursion de longueur* 1 *du mouvement brownien standard.*

Ainsi, dans le cas d'une loi de reproduction de variance finie σ^2, on a la convergence

$$\frac{\sigma}{\sqrt{n}}T_n \xrightarrow[n\to\infty]{(\mathcal{L})} \mathcal{T}^{br} := \mathcal{T}^{(2\mathbf{e})},$$

au sens de Gromov-Prokhorov (ainsi qu'au sens de Gromov-Hausdorff-Prokhorov). L'arbre $\mathcal{T}^{br}$ est appelé arbre continu brownien. Notons que différentes conventions existent pour le choix de la constante apparaissant devant l'excursion du mouvement brownien standard. Celle qui apparaît ici est la même que dans les articles d'Aldous [5, 6, 7].

Dans le second cas, la limite d'échelle que l'on obtient est l'arbre stable d'indice α. Rappelons les éléments essentiels de sa définition, due à Duquesne et Le Gall [32] (voir également [45]). Soit $\alpha \in (1, 2]$. On note $X^{(\alpha)}$ un processus de Lévy stable d'indice α, à sauts positifs, de transformée de Laplace

$$\mathbb{E}[e^{-\lambda X_t^{(\alpha)}}] = e^{t\lambda^\alpha}.$$

Pour tout $t > 0$, on considère le processus "renversé" $\hat{X}^{(\alpha,t)}$ défini par

$$\hat{X}_s^{(\alpha,t)} = \begin{cases} X_t^{(\alpha)} - X_{(t-s)^-}^{(\alpha)} & \text{if } 0 \leq s < t \\ X_t^{(\alpha)} & \text{if } s = t, \end{cases}$$

et on note $\hat{S}_s^{(\alpha,t)} = \sup_{0\leq r\leq s} \hat{X}_r^{(\alpha,t)}$ pour tout $s \in [0, t]$. On note finalement $H^{(\alpha)}$ le processus tel que $H_0^{(\alpha)} = 0$, et pour tout $t > 0$, $H_t^{(\alpha)}$ est le temps local au niveau 0 du processus $\hat{S}^{(\alpha,t)} - \hat{X}^{(\alpha,t)}$, au temps t, avec la normalisation suivante :

$$H_t^{(\alpha)} = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_0^t \mathbf{1}_{\{\hat{S}_s^{(\alpha,t)} - \hat{X}_s^{(\alpha,t)} < \varepsilon\}} ds \qquad \text{p.s.}$$

Ce processus admet une modification continue, qui est celle que l'on considère par la suite (voir [32, Section 1.2]). L'arbre stable d'indice α est l'arbre

$$\mathcal{T}^\alpha := \mathcal{T}^{(H^{[\alpha]})},$$

avec $H^{[\alpha]}$ l'excursion de longueur 1 du processus $H^{(\alpha)}$.

Théorème (Duquesne [31])**.** *Dans le cas 2, il existe une suite de la forme $a_n = n^{1/\alpha}L(n)$, avec L une fonction à variation lente, telle que l'on ait la convergence*

$$\left(\frac{a_n}{n}C_{T_n}(2nt)\right)_{0\leq t\leq 1} \xrightarrow[n\to\infty]{(\mathcal{L})} \left(H_t^{[\alpha]}\right)_{0\leq t\leq 1}.$$

Ce résultat montre en particulier la convergence

$$\frac{a_n}{n}T_n \xrightarrow[n\to\infty]{(\mathcal{L})} \mathcal{T}^\alpha \tag{1.1}$$

au sens de Gromov-Prokhorov (ainsi qu'au sens de Gromov-Hausdorff-Prokhorov), dans le cas où ν appartient au bassin d'attraction de la loi stable d'indice α. Notons que pour $\alpha = 2$, ce résultat est cohérent avec celui du premier cas. En effet, l'arbre stable d'indice $\alpha = 2$ correspond à un facteur d'échelle près à l'arbre brownien : on a $\sqrt{2} \cdot \mathcal{T}^2 = \mathcal{T}^{br}$.

1.2 Fragmentations d'arbres aléatoires

De manière générale, on appelle *fragmentation* un processus décrivant la dislocation progressive d'un objet.

1.2.1 Processus de coupe d'arbres discrets

Dans le cadre discret, les fragmentations que nous considérons se font en temps discret, et sont définies de manière assez simple, par des suppressions successives d'arêtes de l'arbre initial. Ce type de processus a été introduit par Meir et Moon [48], avec le modèle suivant : partant d'un arbre $T \in \mathbf{T}$, à chaque étape, on efface une arête de T choisie uniformément. On appelle *fragmentation aux arêtes* de T ce processus.

Nos travaux portent également sur un autre modèle, que l'on appelle *fragmentation aux nœuds*, défini de la manière suivante. Soit $T \in \mathbf{T}$. À chaque étape :

— on marque un sommet de T choisi avec probabilité proportionnelle à son degré sortant (c'est-à-dire son nombre de fils),

— on supprime toutes les arêtes allant de v vers ses fils, avec v le sommet marqué.

Comme pour la fragmentation aux arêtes, on répète cette opération jusqu'à ce que toutes les arêtes aient disparu.

Dans le cas où T est un arbre aléatoire, ces deux processus sont définis de la même manière, en conditionnant préalablement par rapport à T.

1.2.2 Fragmentations autosimilaires de l'arbre stable et de l'arbre brownien

Les analogues continus de ces deux fragmentations présentent une propriété d'*auto-similarité*. Commençons par expliquer plus précisément ce que signifie cette propriété, dans le cas de processus de fragmentation assez généraux.

Considérons un objet de masse 1 qui se disloque au cours du temps. On peut représenter un tel processus par la suite décroissante des masses des fragments créés au fur et à mesure. On appelle donc *fragmentation* un processus càdlàg à valeurs dans l'ensemble

$$\mathcal{S} = \{\underline{s} = (s_1, s_2, \ldots) : s_1 \geq s_2 \geq \ldots \geq 0, s_i \leq 1 \ \forall i \in \mathbb{N}\}.$$

Les fragmentations que nous étudierons sont des fragmentations auto-similaires : deux fragments séparés, à partir d'un instant t, se désagrègent de manière indépendante, à une vitesse dépendant de leur masse. Plus précisément, une fragmentation auto-similaire d'indice $\beta \in \mathbb{R}$ est un processus de Markov $(F(t), t \geq 0)$ à valeurs dans $\mathcal{S}$, continu en probabilité, tel que $F(0) = (1, 0, 0, \ldots)$, qui vérifie la propriété suivante : sachant $F(t) = (s_1, s_2, \ldots)$, $F(t+t')$ a même loi que le réarrangement décroissant des suites $s_i F^{(i)}(s_i^\beta t)$, avec $F^{(1)}, F^{(2)}, \ldots$ des processus indépendants suivant la même loi que F.

Ces fragmentations ont été étudiées en détail par Bertoin [15]. L'un de ses résultats essentiels est le suivant : la loi d'un tel processus est caractérisée par un triplet (β, c, ν), avec β l'indice d'auto-similarité, $c \geq 0$ un "coefficient d'érosion", et ν une mesure σ-finie sur $\mathcal{S}$, pour laquelle la fonction $\underline{s} \to (1 - s_1)$ est intégrable, et telle que $\nu(1, 0, 0, \ldots) = 0$. Cette mesure, appelée mesure de dislocation, décrit la loi de la suite des masses des fragments qui se créent lors d'une dislocation. Cette propriété permet de caractériser de manière relativement simple la loi d'une fragmentation auto-similaire.

On va s'intéresser à deux mécanismes de fragmentation sur les arbres aléatoires : une fragmentation de l'arbre brownien le long de son squelette, et une fragmentation aux nœuds de l'arbre stable. Ces processus ont été étudiés, respectivement, par Aldous et Pitman [9] et par Miermont [50]. Pour définir ces fragmentations, on travaille conditionnellement à l'arbre aléatoire considéré ($\mathcal{T}^{br}$ ou $\mathcal{T}^\alpha$), et pour un représentant fixé de sa classe d'équivalence.

On commence par introduire deux nouvelles mesures sur nos arbres aléatoires. La première, appelée *mesure de longueur*, peut être définie pour n'importe quel arbre continu $(\mathcal{T}, d, \mu, \rho)$: c'est la mesure σ-finie λ telle que

$$\lambda([\![x, y]\!]) = d(x, y)$$

pour tous points $x, y \in \mathcal{T}$. On peut la voir comme la mesure uniforme sur le squelette de $\mathcal{T}$.

La seconde concerne plus spécifiquement l'arbre stable ; elle est répartie sur les points de branchement de $\mathcal{T}^\alpha$, et pondérée par leur "taille". Plus précisément, pour un arbre continu $\mathcal{T}$, on appelle multiplicité d'un point $x \in \mathcal{T}$ le nombre de composantes connexes de $\mathcal{T} \setminus \{x\}$, noté $\mathrm{mult}(x)$; on note $\mathcal{B}(\mathcal{T})$ l'ensemble des points de multiplicité supérieure ou égale à 3, appelés points de branchement de $\mathcal{T}$. Presque sûrement, dans l'arbre brownien, tous les points de branchement sont de multiplicité 3, tandis que dans l'arbre stable d'indice $\alpha \in (1,2)$, tous les points de branchement sont de multiplicité infinie. Pour mesurer la taille d'un point de branchement b de $\mathcal{T}^\alpha$, on utilise donc la quantité suivante :

$$L(b) = \lim_{\varepsilon \to 0} \varepsilon^{-1} \mu \left(x \in \mathcal{T}^\alpha : b \in [\![\rho, x]\!] \text{ et } d(b,x) \leq \varepsilon \right).$$

Presque sûrement, pour tout $b \in \mathcal{B}(\mathcal{T}^\alpha)$, avec $\alpha \in (1,2)$, cette quantité est bien définie et strictement positive (voir [33, 50]).

Fragmentation le long du squelette de l'arbre brownien : On note $(t_i, \chi_i)_{i \in \mathbb{N}}$ les atomes d'un processus ponctuel de Poisson $\mathcal{N}_{br}$ de mesure d'intensité $dt \otimes \lambda(dx)$ sur $\mathbb{R}_+ \times \mathcal{T}^{br}$. On fragmente l'arbre $\mathcal{T}^{br}$ en supprimant le point χ_i à l'instant t_i, pour tout $i \in \mathbb{N}$. Ainsi, l'arbre fragmenté à l'instant t est

$$\mathcal{T}^{br} \setminus \{\chi_i : t_i \leq t\}.$$

Le processus $(F_{br}(t))_{t \geq 0}$ associé, à l'instant t, est égal à la suite décroissante des masses des composantes connexes de $\mathcal{T}^{br} \setminus \{\chi_i : t_i \leq t\}$. L'introduction de cette fragmentation, étudiée par Aldous et Pitman dans [9], a été motivée par l'étude d'un autre processus appelé coalescent additif standard, auquel elle est reliée par un changement de temps. Les résultats de [9] montrent en particulier que F_{br} est une fragmentation autosimilaire d'indice $1/2$, de coefficient d'érosion $c = 0$, et de mesure de dislocation ν_{br} définie par

$$\nu_{br}\left(\underline{s} \in \mathcal{S} : s_1 \in dx\right) = \frac{dx}{2\pi x^3 (1-x)^3} \qquad \forall x \in [1/2, 1[\qquad \text{et} \qquad \nu_{br}\left(\underline{s} \in \mathcal{S} : s_1 + s_2 < 1\right) = 0.$$

Dans le cas de l'arbre stable, une telle fragmentation (homogène sur le squelette) n'est pas autosimilaire, en raison de la présence de points de branchement de degré infini. En effet, la "taille" de ces points de branchement n'est pas affectée par la fragmentation, car les coupes n'apparaissent jamais au niveau d'un tel point. C'est ce qui nous amène à considérer plutôt une fragmentation aux points de branchement. Notons que des fragmentations plus générales, sur des arbres de Lévy, ont par exemple été étudiées par Abraham, Delmas et Voisin [3].

Fragmentation aux nœuds de l'arbre stable : Soit $\alpha \in (1,2)$. On note $(t_i, \chi_i)_{i \in \mathbb{N}}$ les atomes d'un processus ponctuel de Poisson $\mathcal{N}_\alpha$ de mesure d'intensité $dt \otimes \sum_{b \in \mathcal{B}(\mathcal{T}^\alpha)} L(b) \delta_b(dx)$ sur $\mathbb{R}_+ \times \mathcal{T}^\alpha$. Comme précédemment, on fragmente l'arbre $\mathcal{T}^\alpha$ en le "coupant" au point χ_i à l'instant t_i, pour tout $i \in \mathbb{N}$. Le processus associé $(F_\alpha(t))_{t \geq 0}$ est égal, à l'instant t, à la suite décroissante des masses des composantes connexes de $\mathcal{T}^\alpha \setminus \{\chi_i : t_i \leq t\}$. Ce processus, étudié par Miermont [50], est également une fragmentation autosimilaire, pour laquelle l'indice d'autosimilarité est égal à $1/\alpha$, le coefficient d'érosion est nul, et la mesure de dislocation ν_α est caractérisée par

$$\nu_\alpha(\Phi) = D_\alpha \mathbb{E}\left[T_1 \Phi(T_1^{-1} \Delta T_{[0,1]})\right]$$

pour toute fonction mesurable Φ, avec D_α une constante, $(T_s)_{0 \leq s \leq 1}$ un subordinateur stable de transformée de Laplace

$$\mathbb{E}[e^{-\lambda T_s}] = e^{-\lambda^{1/\alpha} s},$$

et $\Delta T_{[0,1]}$ la suite décroissante des sauts du processus T.

Ces deux fragmentations présentent une propriété intéressante de dualité avec les fragmentations "par la hauteur" de l'arbre brownien et de l'arbre stable, respectivement étudiées par Bertoin [14] et Miermont [49]. En effet, ces dernières sont également des fragmentations auto-similaires, d'indices respectifs $-1/2$ et $1/\alpha - 1$, de coefficient d'érosion nul et de mesures de dislocation ν_{br} et ν_α.

Dans la suite, afin de rendre les notations plus homogènes, les résultats concernant la fragmentation de l'arbre brownien seront parfois énoncés pour une fragmentation analogue sur l'arbre stable $\mathcal{T}^2$, dont les points de coupe sont donnés par un processus de Poisson $\mathcal{N}_2$ d'intensité $dt \times 2\lambda_2(dx)$ (avec λ_2 la mesure de longueur sur $\mathcal{T}^2$). Notons que cette fragmentation est "accélérée" par rapport à celle de l'arbre brownien, afin de conserver une identité en loi qui apparaîtra dans la partie suivante (Théorème 1). Toutefois, les résultats sur $(\mathcal{T}^2, \mathcal{N}_2)$ peuvent toujours se traduire de manière naturelle sur $(\mathcal{T}^{br}, \mathcal{N}_{br})$.

1.3 Arbre des coupes associé à une fragmentation

Outre leurs définitions analogues, les premiers liens entre les modèles de fragmentation discrets et continus apparaissent avec les travaux de Janson [40]. Ce dernier mène une étude asymptotique du nombre de coupes nécessaires pour isoler un sommet donné, pour la fragmentation aux arêtes de T_n, lorsque n tend vers l'infini. (Rappelons que T_n est un arbre de Galton-Watson de loi de reproduction ν, conditionné à avoir n arêtes.) Plus précisément, son résultat peut s'énoncer de la manière suivante. Soit ξ_n un sommet uniforme de T_n, et N_n le nombre de coupes se produisant dans la composante connexe de ξ_n (dans l'arbre fragmenté). Si la loi de reproduction de T_n est de variance finie σ^2, alors on a la convergence en loi

$$\frac{1}{\sigma\sqrt{n}} N_n \xrightarrow[n\to\infty]{(\mathcal{L})} \Delta,$$

avec Δ une variable aléatoire suivant la loi de Rayleigh, de densité $xe^{-x^2/2}$ sur $\mathbb{R}_+$; or cette loi est aussi celle de la distance entre la racine et une feuille uniforme de l'arbre brownien. Un article récent d'Addario-Berry, Broutin et Holmgren [4] donne une preuve simplifiée de cette convergence, établissant un lien plus concret avec l'arbre brownien. Leur résultat a été généralisé par Abraham et Delmas [2] au cas de la fragmentation aux nœuds, pour des arbres de Galton-Watson convergeant vers un arbre stable.

L'arbre des coupes associé à une fragmentation, introduit par Bertoin et Miermont dans [19], décrit la généalogie des différentes composantes connexes de l'arbre fragmenté. En particulier, il code de manière naturelle les "nombres de coupes" nécessaires pour isoler plusieurs sommets. L'étude asymptotique de cet objet permet donc de généraliser le résultat ci-dessus.

Nous détaillons ici la construction de l'arbre des coupes dans les cas discret et continu, ainsi que les résultats de convergence qui font le lien entre ces modèles. Les résultats concernant la fragmentation aux arêtes et le cas brownien sont dus à Bertoin et Miermont [19] ; ceux qui portent sur la fragmentation aux nœuds et le cas stable sont l'objet du Chapitre 1.

1.3.1 Cas discret

Soit $T \in \mathbb{T}$ un arbre discret à n arêtes. On note $e_1, \ldots, e_n$ les arêtes de T, numérotées par exemple dans l'ordre d'un parcours en profondeur. On considère l'une des fragmentations de T définies dans la Section 1.2.1, que l'on décrit ici par le N-uplet $\mathcal{F} = (E_1, \ldots, E_N)$, avec N le nombre d'étapes, et pour tout $r \in \{1, \ldots, N\}$, E_r l'ensemble des arêtes dont la suppression a lieu à l'étape r. Notons que N, $\mathcal{F}$, et éventuellement T, sont des variables aléatoires ; toute la construction suivante se fait conditionnellement à leurs valeurs. L'arbre des coupes associé à la fragmentation $\mathcal{F}$, noté $\mathrm{Cut}(T, \mathcal{F})$, est un arbre discret enraciné qui représente la généalogie des différentes composantes connexes créées au cours du processus.

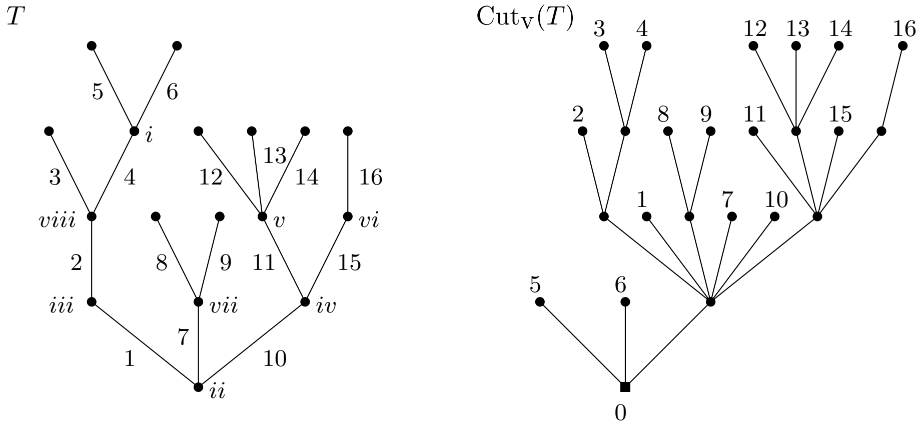

FIGURE 1.2 – Un exemple de construction de l'arbre des coupes, dans le cas d'une fragmentation aux nœuds. Pour l'arbre initial T, les numéros des arêtes sont indiqués en chiffres arabes, et l'ordre de suppression des sommets en chiffres romains. Pour l'arbre des coupes associés, seuls les numéros des feuilles sont indiqués, afin d'alléger la figure ; les nœuds internes sont naturellement étiquetés par l'ensemble des numéros des feuilles qui font partie de leurs descendants. On utilise la notation 0 pour la racine.

Commençons par introduire quelques notations qui permettent de décrire l'ensemble de ses sommets. Pour tout $r \in \{0, \dots, N\}$, on note $D_r = \{i \in \{1, \dots, n\} : e_i \in \bigcup_{r' \leq r} E_{r'}\}$, et on définit une relation d'équivalence $\sim_r$ sur $\{1, \dots, n\} \setminus D_r$ de la manière suivante : on note $j \sim_r j'$ si e_j et $e_{j'}$ appartiennent à la même composante connexe de l'arbre fragmenté à l'issue de l'étape r (autrement dit, s'ils appartiennent au même arbre dans la forêt obtenue en supprimant les arêtes de T appartenant à $\bigcup_{r' \leq r} E_{r'}$). L'ensemble des nœuds internes de $\mathrm{Cut}(T, \mathcal{F})$ est la famille des classes d'équivalence des relations $\sim_0, \dots, \sim_N$, sans répétition. La racine correspond au "bloc" initial $\{1, \dots, n\}$. On y ajoute des feuilles portant les numéros $1, \dots, n$. Notons que pour tout i, la feuille i est distincte du nœud interne $\{i\}$, s'il existe.

La construction de $\mathrm{Cut}(T, \mathcal{F})$ se fait maintenant de manière récursive. On va décrire la r-ème étape de cette construction, pour $r \in \{1, \dots, N\}$. On note B la classe d'équivalence de $\sim_{r-1}$ correspondant à la composante connexe de l'arbre fragmenté à laquelle appartiennent les arêtes de E_r. Remarquons que notre hypothèse sur la fragmentation assure que ces arêtes appartiennent bien à la même composante. Suite à la suppression de ces arêtes, le bloc B se scinde en k'_r parties $B_1, \dots, B_{k'_r}$, qui sont des classes d'équivalence de $\sim_r$. On crée k'_r arêtes allant de B aux nœuds $B_1, \dots, B_{k'_r}$, et $k_r = \#E_r$ arêtes entre B et les feuilles i, pour i tel que $e_i \in E_r$. La figure 1.2 donne un exemple de cette construction.

Notons que $\mathrm{Cut}(T, \mathcal{F})$ est un arbre enraciné, mais n'a pas de structure plane naturelle. Par ailleurs, comme annoncé plus haut, les distances dans l'arbre des coupes correspondent à des "nombres de coupes" dans la fragmentation - plus précisément, la distance de graphe entre la racine et la feuille i, dans $\mathrm{Cut}(T, \mathcal{F})$, correspond au nombre de coupes ayant lieu dans la composante connexe de l'arête i, jusqu'au moment où elle est supprimée.

Pour alléger les notations, on note par la suite $\mathrm{Cut}_{\mathrm{e}}(T)$ l'arbre des coupes associé à une fragmentation aux arêtes de T, et $\mathrm{Cut}_{\mathrm{v}}(T)$ l'arbre des coupes associé à une fragmentation aux nœuds.

1.3.2 Cas continu

Soit $(\mathcal{T}, d, \mu, \rho)$ arbre continu, que l'on fragmente par un processus ponctuel $\mathcal{N}$ sur $\mathbb{R}_+ \times \mathcal{T}$, comme au paragraphe 1.2.2. On note $(t_i, \chi_i)_{i \in \mathbb{N}}$ les atomes de $\mathcal{N}$. Pour tous $x \in \mathcal{T}$, $t \geq 0$, on note $\mathcal{T}_x(t)$ la composante connexe de $\mathcal{T} \setminus \{\chi_i : t_i \leq t\}$ contenant x. Notre but est de construire un arbre $\mathrm{Cut}(\mathcal{T}, \mathcal{N})$ ayant, grossièrement, les propriétés suivantes :

— les feuilles de $\mathrm{Cut}(\mathcal{T},\mathcal{N})$ (excepté la racine) “correspondent” aux points de $\mathcal{T}$,
— pour tout $x \in \mathcal{T}$, la distance entre la racine de $\mathrm{Cut}(\mathcal{T},\mathcal{N})$ et le point correspondant à x dans $\mathrm{Cut}(\mathcal{T},\mathcal{N})$ est un analogue continu du “nombre de coupes” se produisant dans $\mathcal{T}_x(t)$, pour $t \geq 0$.

Pour cela, on définit une nouvelle “distance” sur $\mathcal{T} \sqcup \{0\}$, avec 0 un point supplémentaire qui va correspondre à la racine de l'arbre des coupes. Pour tous $x, y \in \mathcal{T}$, on note

$$\delta(0,0) = 0, \qquad \delta(0,x) = \delta(x,0) = \int_0^\infty \mu(\mathcal{T}_x(t))dt,$$
$$\delta(x,y) = \int_{t(x,y)}^\infty (\mu(\mathcal{T}_x(t)) + \mu(\mathcal{T}_y(t)))dt,$$

avec $t(x,y) := \inf\{t \in \mathbb{R}_+ : \mathcal{T}_x(t) \neq \mathcal{T}_y(t)\}$. Il n'est pas clair a priori que cette quantité soit finie, ni qu'elle définisse une distance sur $\mathcal{T} \sqcup \{0\}$; toutefois, dans le cas où $(\mathcal{T},\mathcal{N}) = (\mathcal{T}^\alpha,\mathcal{N}_\alpha)$ avec $\alpha \in (1,2]$, elle s'interprète de manière naturelle comme un changement de temps liant la fragmentation étudiée à la fragmentation par la hauteur de $\mathcal{T}^\alpha$; elle peut également être vue comme l'analogue de la distance de graphe sur l'arbre discret.

Dans le cas brownien, Bertoin et Miermont [19] définissent l'arbre des coupes de la manière suivante. Soit $(\xi_i)_{i\geq 1}$ une suite de points i.i.d. de $\mathcal{T}^{br}$, de loi μ, et $\xi_0 = 0$. Pour tout $k \geq 1$, δ définit une distance sur $\{\xi_i, 0 \leq i \leq k\}$, et l'espace métrique $\mathcal{R}(k) := (\{\xi_i, 0 \leq i \leq k\}, \delta)$ peut être représenté comme un arbre à k feuilles, enraciné en 0. En outre, les arbres $\mathcal{R}(k)$ vérifient une propriété de consistance (la loi du sous-arbre engendré par la racine et k feuilles distinctes uniformes de $\mathcal{R}(k')$, pour $k' \geq k$, est la même que celle de $\mathcal{R}(k)$), et la propriété suivante :

$$\min_{1\leq j\leq k} \delta(\xi_0,\xi_j) \xrightarrow[k\to\infty]{(\mathcal{L})} 0.$$

Ces caractéristiques assurent l'existence d'un arbre continu aléatoire $\mathrm{Cut}(\mathcal{T}^{br},\mathcal{N}_{br})$ tel que, conditionnellement à $(\mathcal{T}^{br},\mathcal{N}_{br})$, pour tout k, la loi du sous-arbre engendré par la racine et k feuilles indépendantes uniformes de $\mathrm{Cut}(\mathcal{T}^{br},\mathcal{N}_{br})$ est la même que celle de $\mathcal{R}(k)$. En outre, Bertoin et Miermont démontrent l'égalité en loi

$$\mathrm{Cut}(\mathcal{T}^{br},\mathcal{N}_{br}) \overset{(\mathcal{L})}{=} \mathcal{T}^{br}.$$

Dans le cas de l'arbre stable d'indice $\alpha \in (1,2)$, on démontre que l'arbre des coupes $\mathrm{Cut}(\mathcal{T}^\alpha,\mathcal{N}_\alpha)$ peut être défini suivant le même procédé, et l'on obtient un résultat analogue :

Théorème 1. *La loi de l'arbre des coupes* $\mathrm{Cut}(\mathcal{T}^\alpha,\mathcal{N}_\alpha)$ *est la même que celle de* $\mathcal{T}^\alpha$*.*

Notons que la définition du processus de coupes sur $\mathcal{T}^2$ assure que ce résultat reste valable pour $\alpha = 2$. La démonstration de ce résultat repose essentiellement sur le lien entre la fragmentation aux nœuds de l'arbre stable et la fragmentation par la hauteur. Plus précisément, le fait que ces deux fragmentations aient des caractéristiques similaires, comme indiqué à la fin de la Section 1.2.2, permet de montrer que leurs lois sont les mêmes à un changement de temps près ; on en déduit l'égalité en loi

$$(\delta(\xi_i,\xi_j))_{i,j\geq 0} \overset{(\mathcal{L})}{=} (d_{\mathcal{T}^\alpha}(\xi_{i+1},\xi_{j+1}))_{i,j\geq 0},$$

dont la construction de l'arbre des coupes et le Théorème 1 sont des conséquences directes.

On complète ces résultats en montrant qu'il est possible de définir la correspondance souhaitée entre les points de l'arbre initial et les feuilles de l'arbre des coupes, à condition de faire jouer un rôle particulier aux points de coupe. En effet, pour tout $i \in \mathbb{N}$, le point de coupe χ_i n'appartient plus à l'arbre fragmenté après l'instant t_i ; mais il peut être vu comme une feuille des différents sous-arbres séparés à l'instant t_i, comme indiqué sur la Figure 1.3 (à condition de prendre l'adhérence de ces composantes, ce qui est naturel pour conserver des arbres compacts).

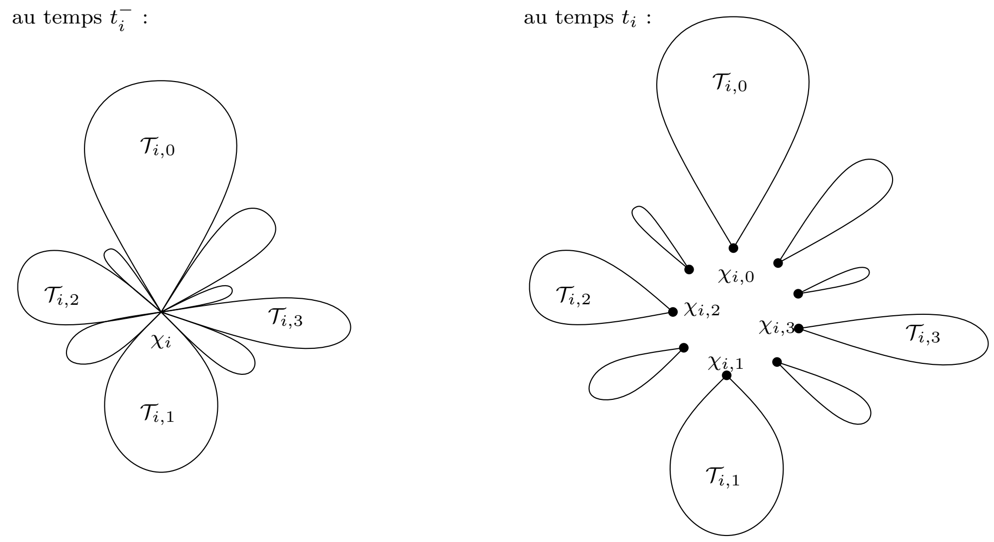

FIGURE 1.3 – À partir de l'instant t_i, le point de coupe χ_i correspond à une feuille des différentes composantes connexes $(\mathcal{T}_{i,j}, j \in \mathbb{N})$ créés par sa suppression. On peut voir ces feuilles comme "différents" points $\chi_{i,j}$, $j \in \mathbb{N}$, dont les images seront des feuilles distinctes de l'arbre des coupes. (Les seules composantes représentées sur la figure sont celles incluses dans $\mathcal{T}_{\chi_i}(t_i^-)$, et non toutes celles de l'arbre initial.)

On introduit donc des points $\chi_{i,j}$ correspondant aux "représentants" de χ_i dans chacune de ces composantes (une définition plus rigoureuse est donnée dans la Section 1.2 du Chapitre 2). L'indice j prend ses valeurs dans un ensemble $\mathcal{J}_\alpha$ ayant deux éléments si $\alpha = 2$, et égal à $\mathbb{N}$ sinon, de manière à avoir $\#\mathcal{J}_\alpha = \text{mult}(\chi_i)$ pour tout i. On considère l'ensemble

$$\mathcal{T}_\mathcal{X}^\alpha = \mathcal{T}^\alpha \setminus \{\chi_i : i \in \mathbb{N}\} \sqcup \{\chi_{i,j} : (i,j) \in \mathbb{N} \times \mathcal{J}_\alpha\}.$$

La définition de δ s'étend de manière naturelle à cet ensemble. Avec ces notations, on démontre l'existence d'une application surjective ℓ de $\mathcal{T}_\mathcal{X}^\alpha$ dans $\mathcal{L}(\mathcal{C}^\alpha) \cup \mathcal{B}(\mathcal{C}^\alpha)$ telle que pour tous $x, y \in \mathcal{T}_\mathcal{X}^\alpha$, on a

$$d_{\mathcal{C}^\alpha}(\ell_x, \ell_y) = \delta(x, y),$$

avec $\mathcal{C}^\alpha$ l'arbre des coupes $\text{Cut}(\mathcal{T}^\alpha, \mathcal{N}_\alpha)$. Un énoncé plus précis est donné dans le Chapitre 2 (Proposition 1.4). La construction de cette correspondance ℓ repose sur le fait que l'on peut obtenir l'arbre des coupes $\text{Cut}(\mathcal{T}^\alpha, \mathcal{N}_\alpha)$ en itérant une transformation de $\mathcal{T}^\alpha$ qui donne sa "première branche" ; cette transformation a été étudiée par Addario-Berry, Broutin et Holmgren [4] dans le cas brownien, et par Abraham et Delmas [2] dans le cas stable.

1.3.3 Convergence de l'arbre des coupes de grands arbres de Galton-Watson

Rappelons que pour tout n, T_n désigne un arbre de Galton-Watson de loi de reproduction ν. Le résultat principal de l'article de Bertoin et Miermont [19] concerne la limite d'échelle des arbres des coupes $\text{Cut}_\text{e}(T_n)$:

Théorème (Bertoin, Miermont [19]). *Si ν est de variance finie σ^2, alors on a la convergence jointe*

$$\left(\frac{\sigma}{\sqrt{n}} T_n, \frac{1}{\sigma\sqrt{n}} \text{Cut}_\text{e}(T_n)\right) \xrightarrow[n\to\infty]{} (\mathcal{T}^{br}, \text{Cut}(\mathcal{T}^{br}, \mathcal{N}_{br}))$$

dans $\mathbb{M} \times \mathbb{M}$, pour la topologie produit associée à la topologie de Gromov-Prokhorov.

On démontre un résultat similaire pour la fragmentation aux nœuds de T_n, lorsque la loi de reproduction ν est de variance finie ou appartient au bassin d'attraction de la loi stable d'indice $\alpha \in (1,2)$. Dans ce dernier cas, on a besoin d'une hypothèse technique supplémentaire :

$$\sup_{r\geq 1}\left(\frac{r\mathbb{P}(\hat{Z}=r)}{\mathbb{P}(\hat{Z}>r)}\right)<\infty, \tag{1.2}$$

avec $\hat{Z}$ une variable aléatoire suivant la loi ν biaisee par la taille, c'est-à-dire telle que $\mathbb{P}(\hat{Z} = r) = r\nu(\{r\})$. Cette hypothèse est par exemple vérifiée si $\nu(\{r\})$ est équivalent à $c/r^{\alpha+1}$ lorsque $n\to\infty$, pour une constante $c \in (0,\infty)$. Notre résultat est le suivant :

Théorème 2. *1. Si ν est de variance finie σ^2, alors on a la convergence jointe*

$$\left(\frac{\sigma}{\sqrt{n}}T_n, \frac{1}{\sqrt{n}}\left(\sigma+\frac{1}{\sigma}\right)\mathrm{Cut}_{\mathrm{v}}(T_n)\right)\xrightarrow[n\to\infty]{}(\mathcal{T}^{br},\mathrm{Cut}(\mathcal{T}^{br},\mathcal{N}_{br}))$$

dans $\mathbb{M}\times\mathbb{M}$, pour la topologie produit associée à la topologie de Gromov-Prokhorov.

2. On suppose que ν appartient au bassin d'attraction de la loi stable d'indice α et vérifie la condition (1.2). Soit (a_n) une suite vérifiant la convergence (1.1). On a alors la convergence jointe

$$\left(\frac{a_n}{n}T_n, \frac{a_n}{n}\mathrm{Cut}_{\mathrm{v}}(T_n)\right)\xrightarrow[n\to\infty]{}(\mathcal{T}^{\alpha},\mathrm{Cut}(\mathcal{T}^{\alpha},\mathcal{N}_{\alpha})),$$

dans $\mathbb{M}\times\mathbb{M}$.

Notons que dans le premier cas, le facteur $1/\sigma$ obtenu par Bertoin et Miermont est remplacé, pour la fragmentation aux nœuds, par $\sigma + 1/\sigma$. Informellement, cela vient du fait que pour la fragmentation aux nœuds de T_n, le nombre moyen d'arêtes supprimées à chaque étape est de l'ordre de $\sum_k k\nu(\{k\})\times k = \sigma^2+1$; par conséquent, les suppressions d'arêtes se font σ^2+1 fois plus vite que pour la fragmentation aux arêtes.

1.4 Reconstruction d'un arbre fragmenté à partir de son arbre des coupes

Dans le Chapitre 2, on s'intéresse à la question de la reconstruction d'un arbre à partir de son arbre des coupes, également étudiée par Broutin et Wang [21] dans le cas brownien. Fixons $\alpha \in (1,2]$. Afin d'alléger les notations, on note $\mathcal{T}$ l'arbre stable d'indice α, $\mathcal{N} = \mathcal{N}_\alpha$, et $\mathcal{C} = \mathrm{Cut}(\mathcal{T},\mathcal{N})$. Remarquons tout d'abord que la connaissance de $\mathcal{C}$ seul ne suffit pas à retrouver exactement l'arbre $\mathcal{T}$. Afin de mieux comprendre de quelle information supplémentaire on a besoin, on revient sur les correspondances entre les points et composantes connexes de l'arbre fragmenté et leurs "images" dans l'arbre des coupes.

Soit $i\in\mathbb{N}$. Rappelons que les points $(\chi_{i,j}, j\in\mathcal{J}_\alpha)$ peuvent être vus comme les "représentants" de χ_i dans chacune des composantes connexes $\mathcal{T}_{i,j}$ créés par la suppression du point χ_i. La correspondance ℓ associe à chacun de ces points une feuille $\ell_{\chi_{i,j}}$ de $\mathcal{C}$. On note b_i le plus récent ancêtre commun de ces points, c'est-à-dire le point tel que

$$[\![\rho_{\mathcal{C}}, b_i]\!] = \bigcap_{j\in\mathcal{J}_\alpha}[\![\rho_{\mathcal{C}},\ell_{\chi_{i,j}}]\!],$$

avec $\rho_{\mathcal{C}}$ la racine de $\mathcal{C}$. Comme indiqué sur la Figure 1.4, on peut voir chacune des composantes connexes de $\mathcal{C}\setminus\{b_i\}$, excepté celle contenant $\rho_{\mathcal{C}}$, comme l'arbre des coupes de l'un des $\mathcal{T}_{i,j}$, $j\in\mathcal{J}_\alpha$ (pour une restriction adaptée du processus de Poisson $\mathcal{N}$).

Supposons maintenant que l'on sache retrouver les arbres $\mathcal{T}_{i,j}$, $j\in\mathcal{J}_\alpha$. Pour reconstruire $\mathcal{T}_{\chi_i}(t_i^-)$ à partir de ces sous-arbres, il est nécessaire de connaître la position du point χ_i dans chacun d'entre eux. Cette information est codée, dans l'arbre des coupes, par la position des feuilles $\ell_{\chi_{i,j}}$, $j \in \mathcal{J}_\alpha$. On munira donc l'arbre des coupes d'une *décoration* $f_{(\mathcal{T},\mathcal{N})} : \mathcal{B}(\mathcal{C}) \to \mathcal{P}(\mathcal{L}(\mathcal{C}))$ (avec $\mathcal{P}((S))$ l'ensemble des parties d'un ensemble S), définie par la relation

$$f_{(\mathcal{T},\mathcal{N})}(b_i) = \left\{\ell_{\chi_{i,j}} : (i,j) \in \mathbb{N} \times \mathcal{J}_\alpha\right\}.$$

On note $\mathrm{DCut}(\mathcal{T},\mathcal{N})$ le couple $(\mathrm{Cut}(\mathcal{T},\mathcal{N}), f_{(\mathcal{T},\mathcal{N})})$.

De manière générale, l'opération de reconstruction sera définie pour un arbre stable $\mathcal{C}$ d'indice α, muni d'une décoration adéquate f. Pour tout point de branchement $b \in \mathcal{B}(\mathcal{C})$, on note $(\mathcal{C}_b^{(j)}, j \in \mathcal{J}_\alpha)$ les composantes connexes de $\mathcal{C} \setminus \{b\}$, excepté celle contenant la racine. On dit que f est une bonne décoration sur $\mathcal{C}$ si f est une fonction aléatoire de $\mathcal{B}(\mathcal{C})$ dans $\mathcal{P}(\mathcal{L}(\mathcal{C}))$, dont la loi sachant $\mathcal{C}$ est caractérisée par les conditions suivantes :

— les ensembles aléatoires $f(b)$, $b \in \mathcal{B}(\mathcal{C})$ sont indépendants.

— pour tout $b \in \mathcal{B}(\mathcal{C})$, $f(b)$ est de la forme $\left\{f^{(j)}(b), j \in \mathcal{J}_\alpha\right\}$, avec $(f^{(j)}(b), j \in \mathcal{J}_\alpha)$ des variables aléatoires indépendantes, telles que pour tout $j \in \mathbb{N}$, $f^{(j)}(b)$ est une feuille uniforme de $\mathcal{C}_b^{(j)}$.

Pour tous points $b \in \mathcal{B}(\mathcal{C})$ et $l \in \mathcal{L}(\mathcal{C})$ tels que $b \in [\![\rho_\mathcal{C}, l]\!]$, on note $\mathcal{C}_b^l$ la composante connexe de $\mathcal{C} \setminus \{b\}$ contenant l, et $f_l(b)$ l'unique élément de $f(b) \cap \mathcal{C}_b^l$.

On va définir une nouvelle "distance" $\tilde{\delta}$ entre les feuilles de $\mathcal{C}$. Notre construction, qui peut de prime abord paraître artificielle, sera expliquée plus bas. Fixons deux feuilles l, l' de $\mathcal{C}$, distinctes de $\rho_\mathcal{C}$. On note $U = \{\emptyset\} \cup \bigcup_{n\in\mathbb{N}} \{1,2\}^n$. On construit de manière récursive des suites de points $l(u)$, $l'(u)$ et $b(u)$, et une suite de sous-arbres $\mathcal{C}(u)$, indexées par U :

— On prend $l(\emptyset) = l$, $l'(\emptyset) = l'$ et $\mathcal{C}(\emptyset) = \mathcal{C}$.

— Pour tout $u \in U$, on note $b(u) = l(u) \wedge l'(u)$, et

$$\begin{aligned} l(u,1) &= l(u) \qquad & l'(u,1) &= f_{l(u)}(b(u)) \qquad & \mathcal{C}(u,1) &= \mathcal{C}_{b(u)}^{l(u)} \\ l(u,2) &= f_{l'(u)}(b(u)) \qquad & l'(u,2) &= l'(u) \qquad & \mathcal{C}(u,2) &= \mathcal{C}_{b(u)}^{l'(u)}. \end{aligned}$$

Les premières étapes de cette construction sont illustrées sur la Figure 1.5. Pour tout $\varepsilon > 0$, on note $U_\varepsilon^{l,l'} = \{u \in U : \mu_\mathcal{C}(\mathcal{C}(u)) \geq \varepsilon\}$, et $N^{l,l'}(\varepsilon) = \#U_\varepsilon^{l,l'}$. Ainsi, $N^{l,l'}(\varepsilon)$ correspond au nombre de points de branchement du sous-arbre de $\mathcal{C}$ engendré par la racine et les feuilles $l(u)$, $l'(u)$, pour $u \in U_\varepsilon$.

Revenons brièvement au cas où $(\mathcal{C}, f) = \mathrm{DCut}(\mathcal{T},\mathcal{N})$, afin d'expliquer le lien entre $N^{\ell_x,\ell_y}(\varepsilon)$ et la distance $d_\mathcal{T}(x,y)$. Considérons la fragmentation de $\mathcal{T}$ "au niveau ε", obtenue en ne conservant que les coupes qui interviennent dans des composantes de taille supérieure à ε. La quantité

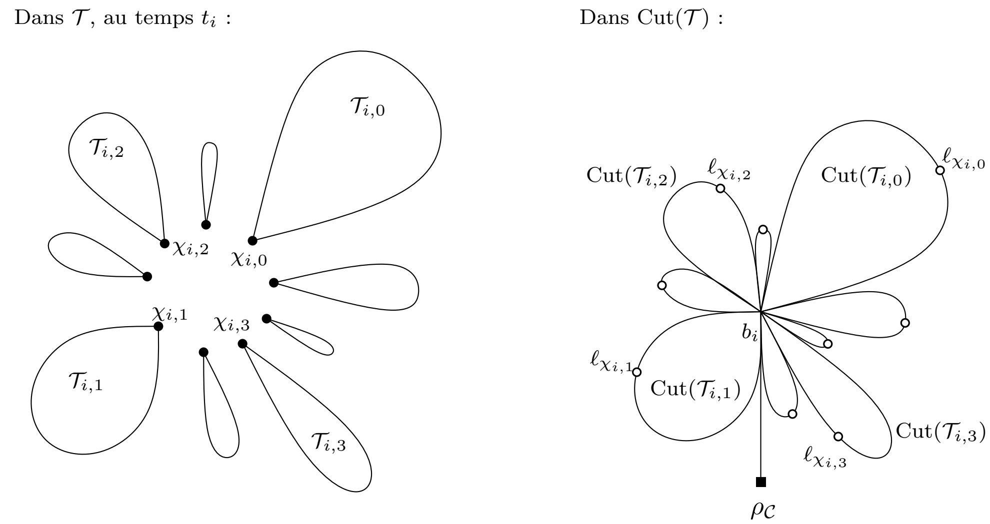

FIGURE 1.4 – Illustration de la correspondance entre les composantes de l'arbre fragmenté et celles de son arbre des coupes.

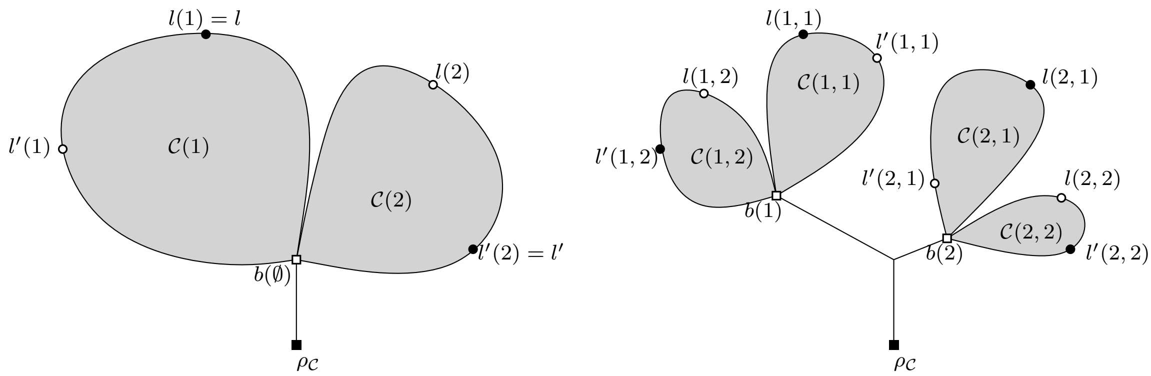

FIGURE 1.5 – Les premières étapes de la construction récursive des suites $l(u)$, $l'(u)$, $b(u)$ et $\mathcal{C}(u)$. À chaque étape, les nouvelles feuilles sont représentées par des sommets blancs ; elle sont obtenues *via* la décoration des points de branchements représentés par des carrés blancs.

$N^{\ell_x,\ell_y}(\varepsilon)$ correspond au nombre de coupes qui se produisent dans le segment $[\![x,y]\!]$ pour cette fragmentation ; il est donc naturel de penser que l'on peut approcher la longueur de $[\![x,y]\!]$ par une limite, à une certaine échelle, de cette quantité.

Dans le cas général, $N^{l,l'}(\varepsilon)$ peut s'interpréter comme l'*énergie* liée à une fragmentation en temps discret de $\mathcal{C}$. Les résultats de Bertoin et Martinez [18] permettent de montrer que si l et l' sont deux feuilles indépendantes uniformes de $\mathcal{C}$, la quantité $\varepsilon^{1-1/\alpha}N^{l,l'}(\varepsilon)$ converge presque sûrement. On montre en outre que, pour une constante

$$C_\alpha = \frac{(\alpha-1)^2\Gamma(1-1/\alpha)}{2\alpha\Gamma(2-2/\alpha)},$$

la loi limite de $C_\alpha \cdot \varepsilon^{1-1/\alpha}N^{l,l'}(\varepsilon)$ est la même que celle de $d_{\mathcal{C}}(l,l')$. Cela nous amène à définir la "distance reconstruite" de la manière suivante :

$$\tilde{\delta}(l,l') = \lim_{\varepsilon\to 0} C_\alpha \cdot \varepsilon^{1-1/\alpha}N^{l,l'}(\varepsilon) \qquad \forall l,l' \in \mathcal{L}(\mathcal{C}) \setminus \{\rho_{\mathcal{C}}\}\,.$$

Si $(l_i)_{i\geq 0}$ est une suite de feuilles indépendantes, uniformes dans $\mathcal{C}$, on montre l'égalité en loi

$$(\tilde{\delta}(l_i,l_j))_{i,j\geq 0} \overset{(\mathcal{L})}{=} (d_{\mathcal{C}}(l_i,l_j)_{i,j\geq 0}.$$

Comme dans la construction de l'arbre des coupes, cette égalité assure l'existence d'un arbre continu aléatoire $\mathrm{Rec}(\mathcal{C},f)$ tel que, conditionnellement à $(\mathcal{C},f)$, pour tout k, la loi du sous-arbre engendré par la racine et k feuilles indépendantes uniformes de $\mathrm{Rec}(\mathcal{C},f)$ est la même que celle de l'arbre $(\{l_i, 0\leq i\leq k\}\,,\tilde{\delta})$. Notons que dans le cas où $(\mathcal{C},f) = \mathrm{DCut}(\mathcal{T},\mathcal{N})$, il faut prendre $l_0 = \ell_{\rho_{\mathcal{T}}}$ s'il l'on souhaite que l'arbre reconstruit soit enraciné au même point. Notre résultat principal est le suivant :

Théorème 3. *La loi de l'arbre reconstruit* $\mathrm{Rec}(\mathcal{C},f)$ *est la même que celle de* $\mathcal{C}$. *En outre, si on a* $(\mathcal{C},f) = \mathrm{DCut}(\mathcal{T},\mathcal{N})$*, alors* $\mathrm{Rec}(\mathcal{C},f) = \mathcal{T}$ *presque sûrement.*

On complète ce résultat en montrant que l'on peut également retrouver un processus ponctuel de Poisson $\mathcal{N}_{(\mathcal{C},f)}$ sur $\mathrm{Rec}(\mathcal{C},f)$ vérifiant les propriétés suivantes :

— Si $(\mathcal{C},f) = \mathrm{DCut}(\mathcal{T},\mathcal{N})$, alors $\mathcal{N}_{(\mathcal{C},f)} = \mathcal{N}$ presque sûrement.

— On a $\mathrm{DCut}(\mathrm{Rec}(\mathcal{C},f),\mathcal{N}_{(\mathcal{C},f)}) = (\mathcal{C},f)$.

On conclut cette section par quelques remarques sur les similitudes et différences entre notre travail et celui de Broutin et Wang [21]. La reconstruction de [21] se fait à partir d'un arbre brownien $\mathcal{C}$ et d'un aléa supplémentaire, la donnée d'une feuille F_b associée à chaque point de branchement $b \in \mathcal{B}(\mathcal{C})$. Cet aléa supplémentaire correspond exactement à notre décoration. En effet, Broutin et Wang démontrent que pour tout $b \in \mathcal{B}(\mathcal{C})$, la composante connexe de $\mathcal{C} \setminus \{b\}$ située au-dessus de b et ne contenant pas F_b contient une unique feuille F'_b telle que, pour tout point $b' \in [\![b, F'_b]\!] \cap \mathcal{B}(\mathcal{C})$, on ait $F_{b'} \in \mathcal{C}_{b'}^{F'_b}$. Avec ces notations, notre décoration est la fonction définie par

$$f(b) = \{F_b, F'_b\} \qquad \forall b \in \mathcal{B}(\mathcal{C}),$$

et la feuille l_0 de $\mathcal{C}$ correspondant à la racine de l'arbre reconstruit est l'unique feuille telle que pour tout $b' \in [\![\rho_{\mathcal{C}}, l_0]\!] \cap \mathcal{B}(\mathcal{C})$, on ait $F_{b'} \in \mathcal{C}_{b'}^{l_0}$. Ainsi, le point de départ de la reconstruction est essentiellement le même dans notre travail et dans [21] ; notre présentation est plus "symétrique", dans la mesure où elle ne fait pas jouer un rôle particulier à l_0, ce qui traduit le fait que la racine de $\mathcal{T}$ ne joue pas un rôle particulier dans sa fragmentation.

Par ailleurs, le résultat que nous obtenons est un peu plus précis que celui de Broutin et Wang. En effet, dans [21], l'arbre reconstruit $\mathrm{Shuff}(\mathcal{C})$ vérifie l'égalité en loi

$$(\mathcal{T}, \mathrm{Cut}(\mathcal{T}, \mathcal{N})) \overset{(\mathcal{L})}{=} (\mathrm{Shuff}(\mathcal{C}), \mathcal{C}),$$

tandis que la correspondance exacte entre points de $\mathcal{T}$ et de $\mathrm{Cut}(\mathcal{T}, \mathcal{N})$ nous permet d'identifier la décoration $f_{(\mathcal{T},\mathcal{N})}$ pour laquelle l'arbre reconstruit est *presque sûrement* égal à l'arbre initial.

1.5 Quelques perspectives

L'ensemble de ces travaux donne une vision assez complète de l'arbre des coupes associé à deux fragmentations auto-similaires, sur l'arbre brownien et l'arbre stable. Dès lors, une question qui émerge naturellement est celle de l'extension de ces constructions à d'autres modèles d'arbres aléatoires, ou à d'autres processus de coupe. On pourrait par exemple s'intéresser à des fragmentations homogènes sur le squelette d'arbre réels assez généraux (par exemple de dimension fractale fixée), ou aux fragmentations d'arbres de Lévy étudiées dans [3]. Toutefois, la démonstration de l'existence de l'arbre des coupes repose pour l'instant sur l'auto-similarité des fragmentations considérées. Il faudra donc avoir recours à de nouvelles méthodes pour traiter ces cas plus généraux.

2 La quadrangulation uniforme infinie du plan ré-enracinée en un point à l'infini

La seconde partie de cette thèse est consacrée à l'étude d'une carte infinie aléatoire, la quadragulation uniforme infinie du plan, ré-enracinée en un point "à l'infini" sur une géodésique.

2.1 Cartes aléatoires finies et infinies

Les cartes aléatoires sont un modèle de géométrie aléatoire qui ont connu un intérêt croissant ces dernières années, notamment motivé par des questions issues de la physique théorique. En combinatoire, les cartes *finies* sont étudiées depuis les travaux de Tutte [57] en 1963. L'aspect probabiliste de leur étude s'est développé plus récemment, en particulier avec les articles d'Angel et Schramm [12, 10], qui introduisent un premier modèle de carte aléatoire *infinie* pouvant être vue comme une triangulation "uniforme" du plan. Des résultats analogues pour les quadrangulations ont ensuite été établis par Krikun [42], menant en particulier à la définition de l'objet central de notre travail, la quadrangulation uniforme infinie du plan (UIPQ). On rappelle ci-dessous les éléments essentiels de cette définition.

Une carte planaire finie peut être définie comme un plongement propre d'un graphe connexe fini dans la sphère de dimension 2, vu à homéomorphisme direct près. Les graphes considérés peuvent avoir des arêtes multiples ou des boucles. On appelle faces d'une carte $\mathfrak{m}$ les composantes connexes du complémentaire de l'union des arêtes de $\mathfrak{m}$. Le degré d'une face est le nombre d'arêtes orientées parcourues quand on en fait le tour dans le sens horaire. On note $\mathcal{M}_f$ l'ensemble des cartes planaires finies *enracinées*, c'est-à-dire munies d'une arête orientée distinguée que l'on appelle la racine.

Pour donner un sens à la notion de carte infinie, on s'intéresse aux limites de cartes planaires dont l'on fait tendre le nombre de faces vers l'infini. Les limites en question peuvent être soit des limites d'échelle, obtenues en multipliant la taille des arêtes par un facteur adapté, dans le même esprit que celui de la section précédente, soit des *limites locales*. Détaillons ce second point de vue, qui est celui que l'on adopte ici.

Soit $\mathfrak{m} \in \mathcal{M}_f$. On définit la boule de rayon r dans $\mathfrak{m}$, notée $B_{\mathfrak{m}}(r)$, comme la carte constituée des arêtes de $\mathfrak{m}$ dont les deux extrémités sont à distance au plus r du sommet racine (c'est-à-dire l'origine de l'arête racine). La topologie locale sur $\mathcal{M}_f$ est la topologie associée à la distance

$$D(\mathfrak{m}, \mathfrak{m}') = (1 + \sup\{r \geq 0 : B_{\mathfrak{m}}(r) = B_{\mathfrak{m}'}(r)\})^{-1} \qquad \forall \mathfrak{m}, \mathfrak{m}' \in \mathcal{M}_f.$$

On note $\mathcal{M}$ le complété de $\mathcal{M}_f$ pour cette topologie.

Pour toute carte finie $\mathfrak{m} \in \mathcal{M}_f$, on dit que $\mathfrak{m}$ est une *quadrangulation* si toutes ses faces sont de degré 4, et une *quadrangulation à bord* si toutes ses faces sont de degré 4, excepté une face distinguée, appelée face externe. Ces définitions s'étendent de manière naturelle aux cartes infinies : une carte $\mathfrak{m} \in \mathcal{M}$ est une quadrangulation infinie (resp. une quadrangulation à bord infini) s'il existe une suite $(\mathfrak{q}_n)_{n\geq 0}$ de quadrangulations finies (resp. de quadrangulations finies à bord) convergeant vers $\mathfrak{m}$, telle que le nombre de faces de $\mathfrak{q}_n$ (resp. le degré de la face extérieure de $\mathfrak{q}_n$) tend vers l'infini. On définit les triangulations de manière analogue, le degré de chaque face étant alors égal à 3.

L'UIPQ est la carte aléatoire infinie $\widetilde{Q}_\infty$ définie comme la limite locale des quadrangulations uniformes à n faces. Notons que la limite d'échelle de telles quadrangulations est également connue : il s'agit de la carte brownienne, introduite par Marckert et Mokkadem [47] (voir également [44] et [51]). Dans le même esprit, Curien et Le Gall ont récemment introduit le plan brownien, qui peut être vu comme limite d'échelle de l'UIPQ [25]. Notons que l'on s'attend à ce qu'un résultat analogue à celui que nous démontrons soit vrai dans ce cadre continu.

Pour conclure cette section, remarquons la nuance entre quadrangulation planaire et quadrangulation du plan. Avec les définitions précédentes, toute quadrangulation infinie peut être vue comme un recollement de quadrilatères qui définit une surface orientable, connexe et séparable (éventuellement avec un bord). On parle de quadrangulation du plan (resp. du demi-plan) si cette surface est homéomorphe à $\mathbb{R}^2$ (resp. $\mathbb{R}_+ \times \mathbb{R}$) ; ce n'est a priori pas le cas pour n'importe quelle limite de quadrangulations planaires, mais on démontre que l'UIPQ est bien une quadrangulation du plan.

2.2 Arbres bien étiquetés et quadrangulations

L'un des outils essentiels de l'étude des quadrangulations finies et infinies est la bijection de Cori-Vauquelin-Schaeffer, introduite dans [24, 56, 23], qui établit une correspondance entre quadrangulations et arbres bien étiquetés. En particulier, les résultats de [22] et [52] montrent que l'UIPQ peut être construite en appliquant cette correspondance à un arbre infini dont les sommets sont munis d'étiquettes positives. Les travaux de Curien, Ménard et Miermont [27] montrent qu'elle peut également être obtenue en appliquant un procédé similaire à un arbre infini étiqueté "uniforme", sans la condition de positivité sur les étiquettes. C'est cette approche que l'on va détailler ici.

Soit $\mathbf{T}$ l'ensemble des arbres discrets enracinés plans (finis ou infinis). Un tel arbre peut naturellement être vu comme un élément de $\mathcal{M}$, l'arête racine étant celle qui va de la racine à son fils le plus à gauche. Un couple (T, l) est un arbre *bien étiqueté* si on a $T \in \mathbf{T}$, et si l est une application qui à chaque sommet de T associe un entier, telle que $|l(\mathrm{u}) - l(\mathrm{v})| \leq 1$ pour toute paire de sommets voisins u, v. On note $\mathbb{T}$ l'ensemble des arbres bien étiquetés.

On s'intéresse en outre à des arbres $T \in \mathbf{T}$ ayant une unique épine dorsale, c'est-à-dire un unique chemin infini $(\mathfrak{s}_i(T))_{i\geq 0}$ tel que $\mathfrak{s}_0(T)$ est le sommet racine de T et pour tout $i \geq 0$,

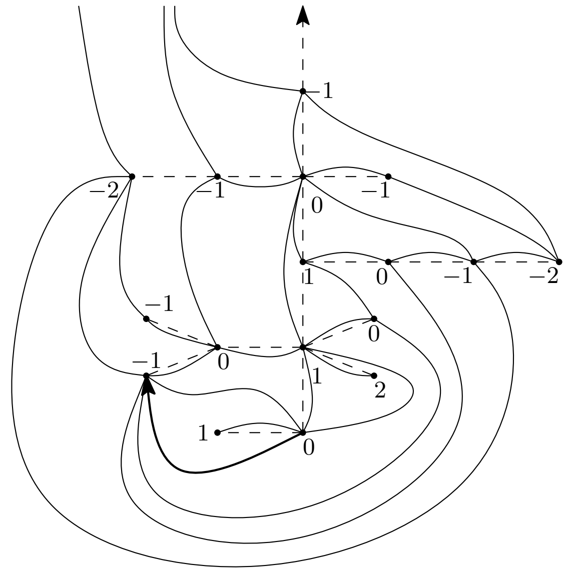

FIGURE 1.6 – La quadrangulation $\Phi(\theta)$ obtenue à partir d'un arbre étiqueté θ. Les arêtes de θ sont représentées en pointillés, et celles de $\Phi(\theta)$ en traits pleins, avec l'arête racine en gras.

$\mathfrak{s}_{i+1}(T)$ est l'un des fils du sommet $\mathfrak{s}_i(T)$. Soit $\mathbf{S}$ l'ensemble de ces arbres. Pour tout $x \in \mathbb{Z}$, on note $\mathbb{S}(x)$ l'ensemble des arbres étiquetés (T, l) tels que $T \in \mathbf{S}$ et la racine de T a pour étiquette x, et

$$\mathbb{S}^*(x) = \left\{ (T, l) \in \mathbb{S}(x) : \inf_{i \geq 0} l(\mathfrak{s}_i(T)) = -\infty \right\}.$$

Pour tout arbre $T \in \mathbf{S}$, pour $i \geq 0$, on note $L_i(T)$ (resp. $R_i(T)$) le sous-arbre formé de $\mathfrak{s}_i(T)$, de ses fils situés à gauche (resp. à droite) de l'épine dorsale et de leurs descendants.

La correspondance définie dans [27] a pour ensemble de départ $\mathbb{S}^*(0)$. Elle est définie de la manière suivante. Soit $\theta = (T, l) \in \mathbb{S}^*(0)$. On appelle *coin* de T tout secteur angulaire défini par deux arêtes de T, n'intersectant aucune autre arête de T. On note c_n, $n \in \mathbb{Z}$ la suite des coins de T, numérotés dans l'ordre donné par un parcours en sens indirect, de manière à ce que c_0 désigne le coin situé à gauche de l'origine de l'arête racine. Pour tout n, on note $l(\mathrm{c}_n)$ l'étiquette du sommet incident à c_n, et on appelle successeur de c_n le premier coin parmi $\mathrm{c}_{n+1}, \mathrm{c}_{n+2}, \ldots$ tel que

$$l(\sigma_\theta(\mathrm{c}_n)) = l(\mathrm{c}_n) - 1.$$

La quadrangulation $\Phi(\theta)$ associée à θ est la carte ayant les mêmes sommets que T, et dont les arêtes sont les chemins reliant chaque coin à son successeur. La figure 1.6 donne un exemple de cette construction. On choisit comme racine l'arête reliant c_0 à son successeur. Notons que l'on peut faire cette construction de telle manière que les arêtes créées ne se croisent pas.

Les travaux de Curien, Ménard et Miermont [27] montrent que pour tout $\theta \in \mathbb{S}(0)$, $\Phi(\theta)$ est une quadrangulation du plan, et que cette correspondance peut être utilisée pour caractériser la loi de l'UIPQ de la manière suivante.

Pour tout $x \in \mathbb{Z}$, on note $\rho_{(x)}$ la loi d'un arbre de Galton–Watson de loi de reproduction $\mathrm{Geom}(1/2)$, tel que la racine a pour étiquette x et que, pour tout sommet v distinct de la racine, l'étiquette de v est uniforme dans $\{\ell - 1, \ell, \ell + 1\}$, avec ℓ l'étiquette de son parent. On considère l'arbre aléatoire $\theta_\infty = (T_\infty, l_\infty) \in \mathbb{S}(0)$ dont la loi est caractérisée par les propriétés suivantes :

— les étiquettes $(S_i(\theta_\infty))_{i\geq 0} := (l_\infty(\mathfrak{s}_i(T_\infty)))_{i\geq 0}$ des sommets de l'épine dorsale sont distribués selon une marche aléatoire à pas indépendants, uniformes dans $\{-1, 0, 1\}$,

— sachant $(S_i(\theta_\infty))_{i\geq 0}$, les arbres $L_i(\theta_\infty)$ et $R_i(\theta_\infty)$, pour $i \geq 0$, sont indépendants dans leur ensemble,

— pour tout $i \geq 0$, sachant $S_i(\theta_\infty)$, $L_i(\theta_\infty)$ et $R_i(\theta_\infty)$ sont des arbres étiquetés de loi $\rho_{(S_i(\theta_\infty))}$.

On peut alors définir l'UIPQ $\widetilde{Q}_\infty$ comme la quadrangulation égale à $\Phi(\theta_\infty)$ avec probabilité 1/2, et obtenue en inversant le sens de l'arête racine de $\Phi(\theta_\infty)$ sinon. En outre, les résultats de [27] montrent que l'on peut inverser cette construction de manière à retrouver θ_∞ à partir de $\widetilde{Q}_\infty$.

2.3 Découpage de l'UIPQ le long de "sa" géodésique infinie

L'une des principales applications la correspondance entre l'UIPQ et l'arbre étiqueté θ_∞ est l'étude des géodésiques infinies issues de la racine. Les étiquettes des sommets de T_∞ peuvent en effet être vues, dans $\widetilde{Q}_\infty$, comme leur "distance à l'infini". Plus précisément, pour tous sommets u, v de $\widetilde{Q}_\infty$, la quantité

$$d_{\widetilde{Q}_\infty}(\mathrm{u},\mathrm{w}) - d_{\widetilde{Q}_\infty}(\mathrm{v},\mathrm{w})$$

est constante en dehors d'un ensemble fini de points w, et égale à $l_\infty(\mathrm{u}) - l_\infty(\mathrm{v})$. (Notons que c'est ce qui permet de retrouver les étiquettes dans la construction inverse de Φ.) En utilisant cette propriété, Curien, Ménard et Miermont démontrent que les géodésiques infinies issues de la racine sont "piégées" entre deux géodésiques particulières, appelées géodésique maximale et géodésique minimale, que l'on peut décrire de façon simple à partir de θ_∞. En outre, ces deux géodésiques ont presque sûrement une infinité de points communs. Toutes les géodésiques infinies issues de la racine sont donc essentiellement similaires.

Dans $\widetilde{Q}_\infty$, la géodésique maximale peut être vue comme la géodésique "la plus à gauche", définie de la manière suivante. Notons e^- et e^+ les deux extrémités de l'arête racine de $\widetilde{Q}_\infty$. D'après la construction de la section précédente, on a

$$l_\infty(e^-) - l_\infty(e^+) = \pm 1.$$

D'après la remarque précédente, si l'on note e_0 et e_1 ces deux extrémités, en imposant

$$l_\infty(\mathrm{e}_0) - l_\infty(\mathrm{e}_1) = 1,$$

l'extrémité e_0 est celle qui est la plus éloignée de l'infini. La géodésique la plus à gauche est définie comme le chemin $(\mathrm{e}_k)_{k\geq 0}$ tel que pour tout $k \geq 1$, e_{k+1} est le premier voisin de e_k dans le sens indirect, après e_{k-1}, tel que

$$l_\infty(\mathrm{e}_k) - l_\infty(\mathrm{e}_{k+1}) = 1.$$

Dans l'arbre étiqueté T_∞, cette géodésique correspond à la chaîne infinie de *coins* définie en prenant $\mathrm{e}_0 = \mathrm{c}_0$ (le coin racine), et pour tout $k \geq 0$, e_k le premier coin d'étiquette $-k$ après e_0, dans le sens indirect. Ainsi, pour tout k, e_{k+1} est le successeur de e_k.

On note $Q_\infty^{(k)}$ la quadrangulation $\widetilde{Q}_\infty$ ré-enracinée en l'arête $(\mathrm{e}_k, \mathrm{e}_{k+1})$. Notre résultat principal est l'identification de la limite locale de $Q_\infty^{(k)}$ lorsque $k \to \infty$. À cette fin, on étudie en réalité les quadrangulations situées de part et d'autre de la géodésique $(\mathrm{e}_k)_{k\geq 0}$, en découpant $\widetilde{Q}_\infty$ le long de cette géodésique infinie, comme illustré sur la Figure 1.7. Soit $\mathrm{Sp}(\widetilde{Q}_\infty)$ la quadrangulation ainsi découpée, et $(\mathrm{e}'_k)_{k\geq 1}$ la géodésique formée des copies des sommets $(\mathrm{e}_k)_{k\geq 1}$. On note $\overrightarrow{Q}_\infty^{(k)}$ la quadrangulation obtenue en ré-enracinant $\mathrm{Sp}(\widetilde{Q}_\infty)$ en $(\mathrm{e}_k, \mathrm{e}_{k+1})$, et $\overleftarrow{Q}_\infty^{(k)}$ la quadrangulation obtenue en ré-enracinant $\mathrm{Sp}(\widetilde{Q}_\infty)$ en $(\mathrm{e}'_k, \mathrm{e}'_{k+1})$.

Avec ces notations, les quadrangulations à bord $\overrightarrow{Q}_\infty^{(k)}$ et $\overleftarrow{Q}_\infty^{(k)}$ peuvent être vues, respectivement, comme les images par Φ des arbres $\theta_\infty^{(k)} = (T_\infty^{(k)}, l_\infty^{(k)})$ et $\theta_\infty^{(-k+1)} = (T_\infty^{(-k+1)}, l_\infty^{(-k+1)})$, avec :

- $T_\infty^{(k)}$ l'arbre T_∞ ré-enraciné en e_k,
- $T_\infty^{(-k+1)}$ l'arbre T_∞ ré-enraciné au coin e_{-k+1} défini comme le dernier coin d'étiquette $-k+1$ avant la racine (toujours dans le sens indirect),

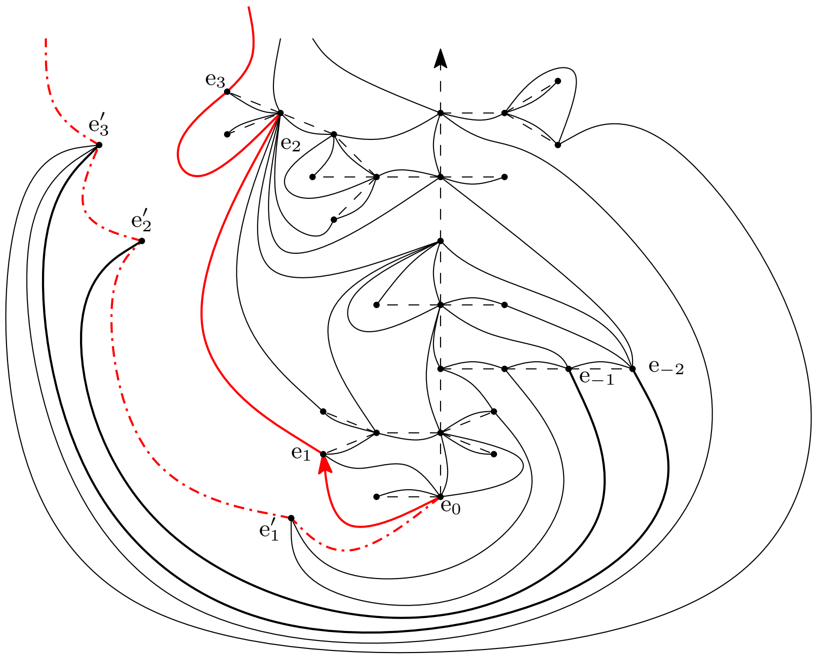

FIGURE 1.7 – La quadrangulation $\mathrm{Sp}(\widetilde{Q}_\infty)$ obtenue en découpant $\widetilde{Q}_\infty$ le long de la géodésique maximale. Les arêtes de l'arbre sous-jacent sont représentées en pointillés, et la géodésique maximale en rouge.

- $l_\infty^{(k)} = l_\infty^{(-k+1)} = l_\infty + k$.

(Notons que $\theta_\infty^{(-k+1)}$ n'appartient pas à $\mathbb{S}^*(0)$ mais à $\mathbb{S}^*(1)$; toutefois, on peut aisément étendre Φ à cet ensemble, avec une convention d'enracinement adaptée.) La première étape de notre travail consiste donc à déterminer la limite en loi du couple $(\theta_\infty^{(k)}, \theta_\infty^{(-k+1)})$. On obtient la convergence jointe

$$(\theta_\infty^{(k)}, \theta_\infty^{(-k+1)}) \xrightarrow[k\to\infty]{(\mathcal{L})} \left(\overrightarrow{\theta_\infty}, \overleftarrow{\theta_\infty}\right) \tag{1.3}$$

pour la topologie locale, avec $\overrightarrow{\theta_\infty}$ et $\overleftarrow{\theta_\infty}$ des arbres étiquetés indépendants à valeurs dans $\mathbb{S}(0)$ et $\mathbb{S}(1)$, dont on détermine les lois explicitement. Les arbres $\overrightarrow{\theta_\infty}$ et $\overleftarrow{\theta_\infty}$ appartiennent, respectivement, aux ensembles

$$\overrightarrow{\mathbb{S}} = \left\{(T,l) \in \mathbb{S}(0) : \min_{n\le -1} l(\mathrm{c}_n(T)) = 1, \lim_{n\to-\infty} l(\mathrm{c}_n(T)) = +\infty \text{ et } \inf_{n\ge 0} l(\mathrm{c}_n(T)) = -\infty\right\}$$

et

$$\overleftarrow{\mathbb{S}} = \left\{(T,l) \in \mathbb{S}(1) : \min_{n\ge 1} l(\mathrm{c}_n(T)) = 2, \lim_{n\to+\infty} l(\mathrm{c}_n(T)) = +\infty \text{ et } \inf_{n\le 0} l(\mathrm{c}_n(T)) = -\infty\right\}.$$

Cela nous amène à étendre la correspondance Φ à des arbres étiquetés appartenant à $\overrightarrow{\mathbb{S}}$ et $\overleftarrow{\mathbb{S}}$. Pour les arbres étiquetés $\theta \in \overrightarrow{\mathbb{S}}$, la construction se fait exactement de la même manière que dans la Section 2.2. Pour $\theta \in \overleftarrow{\mathbb{S}}$, il existe des coins n'ayant pas de successeurs ; on complète donc la construction en ajoutant un bord infini dont les sommets sont étiquetés par les entiers naturels, et servent de successeurs à ces coins. Cette construction est détaillée dans la Section 1.3 du Chapitre 3. Dans les deux cas, la carte obtenue $\Phi(\theta)$ est une quadrangulation du demi-plan à bord géodésique.

Notre résultat principal est le suivant :

Théorème 4. *On note* $\overrightarrow{Q}_\infty = \Phi(\overrightarrow{\theta_\infty})$, $\overleftarrow{Q}_\infty = \Phi(\overleftarrow{\theta_\infty})$, *et* $\overleftrightarrow{Q}_\infty$ *la quadrangulation du plan obtenue en recollant* $\overrightarrow{Q}_\infty$ *et* $\overleftarrow{Q}_\infty$ *le long de leurs frontières respectives, de manière à ce que leurs racines soient identifiées. On a la convergence en loi*

$$Q_\infty^{(k)} \xrightarrow[k\to\infty]{(\mathcal{L})} \overleftrightarrow{Q}_\infty$$

pour la topologie locale.

On obtient ainsi une quadrangulation dans laquelle la racine appartient à une géodésique doublement infinie, dont les "extrémités" correspondent au point à l'infini et au sommet racine de l'UIPQ $\widetilde{Q}_\infty$. Notons que, malgré le fait que les lois de $\overrightarrow{\theta_\infty}$ et $\overleftarrow{\theta_\infty}$ soient symétriques, ce n'est pas le cas pour celles de $\overrightarrow{Q}_\infty$ et $\overleftarrow{Q}_\infty$. En effet, entre deux points du bord de $\overleftarrow{Q_\infty}$, il existe une unique géodésique (égale au bord), tandis qu'on peut observer dans $\overrightarrow{Q}_\infty$ un "faisceau" de géodésiques le long du bord. Cela résulte du choix de la géodésique la plus à gauche dans le ré-enracinement de l'IUPQ.

Une perspective naturelle, dans une étude plus approfondie de la loi de $\overleftrightarrow{Q}_\infty$, serait de déterminer les symétries de sa distribution. Plus précisément, on peut envisager les deux transformations suivantes :

— on commence par ré-enraciner $\overleftrightarrow{Q}_\infty$ le long de sa géodésique "la plus à droite", en un point situé à la même distance de l'infini que sa racine initiale, puis on effectue une symétrie par rapport à la (nouvelle) arête racine ;

— on effectue le même ré-enracinement que dans le cas précédent, puis on inverse le sens de l'arête racine.

Les symétries de la loi de l'UIPQ devraient assurer l'invariance de la loi de $\overleftrightarrow{Q}_\infty$ par la première transformation. La seconde symétrie, qui revient à échanger les rôles de e_0 et du point "à l'infini" de l'UIPQ, est plus délicate à démontrer, et fait l'objet d'une étude en cours.

Chapitre 2

The vertex-cut-tree of Galton-Watson trees converging to a stable tree

Les résultats de ce chapitre sont tirés de l'article [28], publié dans Annals of Applied Probability*.*

1 Introduction and main result

Fragmentations of random trees were first introduced in the work of case of Meir and Moon [48] as a recursive random edge-deletion process on discrete trees. Since then, it has been recognized that fragmentations of discrete and continuous trees appear in several natural contexts: see, for example [17, 30] for a connection with forest fire models, [9, 14] for fragmentations of the Brownian tree [7] and its relation to the additive coalescent, and [3, 49, 50] for fragmentations of the stable tree of index $\alpha \in (1, 2)$ [32]. The fragmentations considered in the two last cases, which arise naturally in the setting of Brownian and stable trees, are self-similar fragmentations as studied by Bertoin [15], whose characteristics are explicitly known.

Several recent articles investigated the question of the asymptotic distribution of the number of cuts needed to isolate a specific vertex, for various classes of random trees. In specific cases, Panholzer [54] showed that the Rayleigh distribution arises naturally as a limit in this context, and Janson [40] showed that this limiting result holds for general Galton–Watson trees with a finite variance offspring distribution, using a method of moments. He also established a connection to the Brownian tree, which is natural since the Rayleigh distribution is the law of the distance between two uniformly chosen vertices in the CRT. Later, Addario-Berry, Broutin and Holmgren [4] provided a different proof giving a more concrete connection to the Brownian tree. Bertoin and Miermont [19] then studied the genealogy of the cutting procedure in itself, which is related to the problem of the isolation of several vertices rather than just the root (certain of these ideas were implicitly present in former papers, including [4, 17]). This allows to code the discrete cutting procedure in terms of a "cut-tree," whose scaling limit is shown to be a Brownian tree, that describes in some sense the genealogy of the Aldous–Pitman fragmentation [9].

Note that the results of [4], by introducing a reversible transformation of the Brownian tree, can be understood as building the "first branch" of the limiting cut-tree, the latter being a kind of iteration *ad libitum* of this transformation. This transformation was extended in [2] in the context of a fragmentation of stable trees. The main goal of the present work is to show that the approach of Bertoin and Miermont [19] can also be adapted to Galton–Watson trees with offspring distribution in the domain of attraction of a non-Gaussian stable law, showing the convergence of the whole discrete cut-tree to a limiting stable tree. This gives in passing a natural definition of the continuum cut-tree for the fragmentation studied in [50].

Let us describe more precisely the result of [19] we are interested in. Consider a sequence of Galton–Watson trees $\mathcal{T}_n$, conditioned to have exactly n edges, with critical offspring distribution having finite variance σ^2. The associated cut-trees $\mathrm{Cut}(\mathcal{T}_n)$ describe the genealogy of the fragments obtained by deleting the edges in a uniform random order. It is well known that the rescaled trees $(\sigma/\sqrt{n}) \cdot \mathcal{T}_n$ converge in distribution to the Brownian tree $\mathcal{T}$; see [7] for the convergence of the associated contour functions, which implies that this convergence holds for the commonly used Gromov-Hausdorff topology, and for the Gromov–Prokhorov topology. In the present work, we will mainly use the latter. Bertoin and Miermont showed that there is in fact the joint convergence

$$\left(\frac{\sigma}{\sqrt{n}}\mathcal{T}_n, \frac{1}{\sigma\sqrt{n}}\,\mathrm{Cut}(\mathcal{T}_n)\right) \xrightarrow[n\to\infty]{(d)} (\mathcal{T}, \mathrm{Cut}(\mathcal{T}))\,,$$

where $\mathrm{Cut}(\mathcal{T})$ is the so-called cut-tree of $\mathcal{T}$. Informally, $\mathrm{Cut}(\mathcal{T})$ describes the genealogy of the fragments obtained by cutting $\mathcal{T}$ at points chosen according to a Poisson point process on its skeleton. Moreover, $\mathrm{Cut}(\mathcal{T})$ has the same law as $\mathcal{T}$.

Our goal is to show an analogue result in the case where the $\mathcal{T}_n$ are Galton–Watson trees with offspring distribution belonging to the domain of attraction of a stable law of index $\alpha \in (1,2)$, and $\mathcal{T}$ is the stable tree of index α. For the stable tree, a self-similar fragmentation arises naturally by splitting at branching points with a rate proportional to their "width," as shown in [50]. This will lead us to modify the edge-deletion mechanism for the discrete trees, so that the rate at which internal vertices are removed increases with their degree. Therefore, we call *edge-fragmentation* the fragmentation studied in [19], and *vertex-fragmentation* our model. Note that more general fragmentations of the stable tree can be constructed by splitting both at branching points and at uniform points of the skeleton, as in [3]. However, these fragmentations are not self-similar (see [50]), and will not be studied here.

In the rest of the Introduction, we will describe our setting more precisely and give the exact definition of the cut-trees, both in the discrete and the continuous cases. This will enable us to state our main results in Section 1.4.

1.1 Vertex-fragmentation of a discrete tree

We begin with some notation. Let $\mathbb{T}$ be the set of all finite plane rooted trees. For every $T \in \mathbb{T}$, we call $E(T)$ the set of edges of T, $V(T)$ the set of vertices of T, and $\rho(T)$ the root-vertex of T. For each vertex $v \in V(T)$, $\deg(v,T)$ denotes the number of children of v in T (or $\deg v$, if this notation is not ambiguous), and for each edge $e \in E(T)$, e^- (resp. e^+) denotes the extremity of e which is closest to (resp. furthest away from) the root.

For any tree T with n edges, we label the vertices of T by $v_0, v_1, \ldots, v_n$, and the edges of T by $e_1, \ldots, e_n$, in the depth-first order. Note that the planar structure of T gives an order on the offspring of each vertex, say "from left to right," hence the depth-first order is well defined. With this notation, we have $v_j = e_j^+$ for all $j \in \{1, \ldots, n\}$.

We let $T \in \mathbb{T}$ be a finite tree with n edges. We consider a discrete-time fragmentation on T, which can be described as follows:

— at each step, we mark a vertex of T at random, in such a way that the probability of marking a given vertex v is proportional to $\deg v$;

— when a vertex v is marked, we delete all the edges e such that $e^- = v$.

Note that the total number of steps N is at most n. To keep track of the genealogy induced by this edge-deletion process, we introduce a new structure called the cut-tree of T, denoted by $\mathrm{Cut}_{\mathrm{v}}(T)$.

For all $r \in \{1, \ldots, N\}$, we let $v(r)$ be the vertex which receives a mark at step r, $E_r = \{e \in E(T) : e^- = v(r)\}$ be the set of the edges which are deleted at step r, $k_r = |E_r|$, and for all $r \in \{0, \ldots, N\}$, $D_r = \{i \in \{1, \ldots, n\} : e_i \in \bigcup_{r' \leq r} E_{r'}\}$. We say that $j \sim_r j'$ if and only if e_j and $e_{j'}$ are still connected in the forest obtained from T by deleting the edges in D_r. Thus $\sim_r$ is an equivalence relation on $\{1, \ldots, n\} \setminus D_r$. The family of the equivalence classes (without

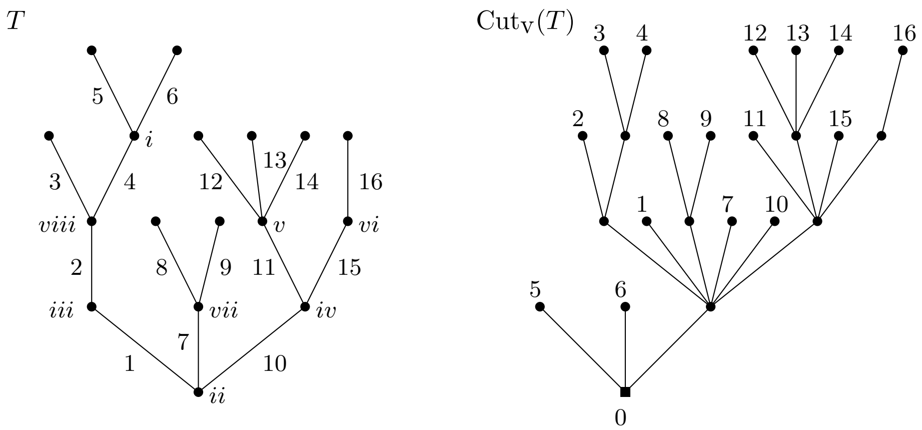

Figure 2.1 – The cut-tree $\mathrm{Cut}_\mathrm{v}(T)$ of a tree T. The order of deletion of the internal vertices of T is indicated in Roman numerals. The correspondence between the edges of T and the leaves of $\mathrm{Cut}_\mathrm{v}(T)$ is indicated in Arabic numerals.

repetition) of the relations $\sim_r$ for $r = 0, \ldots, N$ forms the set of internal nodes of $\mathrm{Cut}_\mathrm{v}(T)$. The initial block $\{1, \ldots, n\}$ is seen as the root, and the leaves of $\mathrm{Cut}_\mathrm{v}(T)$ are given by $1, \ldots, n$. We stress that we distinguish the leaves i and the internal nodes $\{i\}$.

We now build the cut-tree $\mathrm{Cut}_\mathrm{v}(T)$ inductively. At the r-th step, we let B be the equivalence class for $\sim_{r-1}$ containing the indices i such that $e_i \in E_r$. Deleting the edges in E_r splits the block B into k'_r equivalence classes $B_1, \ldots, B_{k'_r}$ for $\sim_r$, with $k'_r \leq k_r + 1$. We draw k'_r edges between B and the sets $B_1, \ldots, B_{k'_r}$, and k_r edges between B and the leaves i such that $e_i \in E_r$. Thus, the graph-distance between the leaf i and the root in $\mathrm{Cut}_\mathrm{v}(T)$ is the number of cuts in the component of T containing the edge e_i before e_i itself is removed. Note that $\mathrm{Cut}_\mathrm{v}(T)$ does not have a natural planar structure, but that the actual embedding does not intervene in our work. Figure 2.1 gives an example of this construction for a tree T with 16 edges.

If T is a random tree, the fragmentation of T and the cut-tree $\mathrm{Cut}_\mathrm{v}(T)$ are defined similarly, by conditioning on T and performing the above construction.

Note that, equivalently, we could mark the edges of T in a uniform random order, and delete all the edges e such that $e^- = e_i^-$, as soon as e_i is marked. The cut-tree $\mathrm{Cut}_\mathrm{v}(T)$ would then be obtained by performing the same construction with $E_r = \{e \in E(T) : e^- = e_{i_r}^-\}$. This procedure sometimes adds "neutral steps," which have no effect on the fragmentation, but this does not change the cut-tree. It will sometimes be more convenient to work with this point of view, for example in Sections 2.1 and 4.

1.2 Fragmentation and cut-tree of the stable tree of index $\alpha \in (1, 2)$

Following Duquesne and Le Gall (see, e.g., [33]), we see stable trees as random rooted $\mathbb{R}$-trees.

Definition 1.1. *A metric space (T, d) is an $\mathbb{R}$-tree if, for every $u, v \in T$,*

— *There exists a unique isometric map $f_{u,v}$ from $[0, d(u, v)]$ into T such that $f_{u,v}(0) = u$ and $f_{u,v}(d(u, v)) = v$.*

— *For any continuous injective map f from $[0, 1]$ into T, such that $f(0) = u$ and $f(1) = v$, we have*

$$f([0, 1]) = f_{u,v}([0, d(u, v)]) := [\![u, v]\!] \,.$$

A rooted $\mathbb{R}$-tree is an $\mathbb{R}$-tree (T, d, ρ) with a distinguished point ρ called the root.

The trees we will work with can be seen as $\mathbb{R}$-trees coded by continuous functions from $[0, 1]$ into $\mathbb{R}_+$, as in [33]. In particular, the stable tree $(\mathcal{T}, d)$ of index α is the $\mathbb{R}$-tree coded by the excursion of length 1 of the height process $H^{(\alpha)}$, defined as follows in [32]. Let $X^{(\alpha)}$ be a stable spectrally positive Lévy process with parameter α, whose normalization will be prescribed in Section 2.2.1. For every $t > 0$, let $\hat{X}^{(\alpha,t)}$ be the process defined by

$$\hat{X}^{(\alpha,t)}_s = \begin{cases} X^{(\alpha)}_t - X^{(\alpha)}_{(t-s)^-} & \text{if } 0 \leq s < t \\ X^{(\alpha)}_t & \text{if } s = t, \end{cases}$$

and write $\hat{S}^{(\alpha,t)}_s = \sup_{0 \leq r \leq s} \hat{X}^{(\alpha,t)}_r$ for all $r \in [0,t]$.

Definition 1.2. *The height process $H^{(\alpha)}$ is the real-valued process such that $H_0 = 0$ and, for every $t > 0$, H_t is the local time at level 0 at time t of the process $\hat{X}^{(\alpha,t)} - \hat{S}^{(\alpha,t)}$.*

The normalization of local time, and the proof of the existence of a continuous modification of this process, are given in [32, Section 1.2]. This definition of $\mathcal{T}$ allows us to introduce the canonical projection $p : [0,1] \to \mathcal{T}$. We endow $\mathcal{T}$ with a probability mass-measure μ defined as the image of the Lebesgue measure on $[0,1]$ under p, and say that the rot of $\mathcal{T}$ is the unique point which has height 0.

For the fragmentation of the stable tree, we will use a process introduced and studied by Miermont in [50], which consists in deleting the nodes of $\mathcal{T}$ in such a way that the fragmentation is self-similar. We first recall that the multiplicity of a point v in an $\mathbb{R}$-tree T can be defined as the number of connected components of $T \setminus \{v\}$. To be consistent with the definitions of Section 1.1, we define the degree of a point as its multiplicity minus 1, and say that a branching point of T is a point v such that $\deg(v,T) \geq 2$. Duquesne and Le Gall have shown in [33, Theorem 4.6] that a.s. the branching points in $\mathcal{T}$ form a countable set, and that these branching points have infinite degree. We let $\mathcal{B}$ denote the set of these branching points. For any $b \in \mathcal{B}$, one can define the local time, or width of b as the almost sure limit

$$L(b) = \lim_{\varepsilon \to 0^+} \varepsilon^{-1} \mu \left\{ v \in \mathcal{T} : b \in [\![\rho, v]\!], d(b,v) < \varepsilon \right\},$$

where ρ is the root of the stable tree $\mathcal{T}$. The existence of this quantity is justified in [50, Proposition 2] (see also [33]).

We can now describe the fragmentation we are interested in. Conditionally on $\mathcal{T}$, we let $(t_i, b_i)_{i \in I}$ be the family (indexed by a countable set I) of the atoms of a Poisson point process with intensity $dt \otimes \sum_{b \in \mathcal{B}} L(b) \delta_b(dv)$ on $\mathbb{R}_+ \times \mathcal{B}$. Seeing these atoms as marks on the branching points of $\mathcal{T}$, we let $\overline{\mathcal{T}}(t) = \mathcal{T} \setminus \{b_i : t_i \leq t\}$.

For every $x \in \mathcal{T}$, we let $\mathcal{T}_x(t)$ be the connected component of $\overline{\mathcal{T}}(t)$ containing x, with the convention that $\mathcal{T}_x(t) = \emptyset$ if $x \notin \overline{\mathcal{T}}(t)$. We also let $\mu_x(t) = \mu(\mathcal{T}_x(t))$. Adding a distinguished point 0 to $\mathcal{T}$, we define a function δ from $(\mathcal{T} \sqcup \{0\})^2$ into $\mathbb{R}_+ \cup \{\infty\}$, such that for all $x, y \in \mathcal{T}$,

$$\delta(0,0) = 0, \qquad \delta(0,x) = \delta(x,0) = \int_0^\infty \mu_x(t) dt,$$
$$\delta(x,y) = \int_{t(x,y)}^\infty (\mu_x(t) + \mu_y(t)) dt,$$

where $t(x,y) := \inf \{ t \in \mathbb{R}_+ : \mathcal{T}_x(t) \neq \mathcal{T}_y(t) \}$ is a.s. finite. We think of δ as our new "distance" in the cut-tree. This definition might seem surprising, but the results of Section 2.1 will show that it provides an analogue of the distance we defined in the discrete case, in terms of number of cuts; as will be explained in Section 3.1, it also has a natural interpretation as a time-change between two fragmentation processes of the stable tree, studied in [49] and [50]. The role of the extra point 0 in our (time-changed) fragmentation will be similar to the role played by the root of $\mathcal{T}$ in the "fragmentation at heights" which will be introduced in Section 3.1.

A first idea would be to build the vertex-cut-tree $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T})$ as a completion of $(\mathcal{T} \sqcup \{0\}, \delta)$. However, making this idea rigorous is difficult, since it is not clear whether δ is a.s. finite, and defines a distance on $\mathcal{T} \sqcup \{0\}$. We will instead use an approach introduced by Aldous, which consists in building a continuous random tree such that the subtrees determined by k randomly chosen leaves have the right distribution. To this end, we use the conditions given by Aldous in [7, Theorem 3].

Set $\xi(0)=0$, and let $(\xi(i))_{i\in\mathbb{N}}$ be an i.i.d. sequence distributed according to μ, conditionally on $\mathcal{T}$. The key argument of our construction is the identity in law

$$(\delta(\xi(i),\xi(j)))_{i,j\geq 0} \overset{(d)}{=} (d(\xi(i+1),\xi(j+1)))_{i,j\geq 0}\,,$$

which will be proven in Section 3.1. In particular, it implies that almost surely, for all $i,j\geq 0$, $\delta(\xi(i),\xi(j))$ is finite, and that δ is a.s. a distance on $\{\xi(i), i\geq 0\}$. This allows us to see the spaces $\mathcal{R}(k) := (\{\xi(i), 0\leq i\leq k\}, \delta)$, for all $k\in\mathbb{N}$, as random rooted trees with k leaves. Using the terminology of Aldous, $(\mathcal{R}(k), k\in\mathbb{N})$ forms a *consistent* family of random rooted trees which satisfies the *leaf-tight condition*:

$$\min_{1\leq j\leq k} \delta(\xi(0),\xi(j)) \xrightarrow[k\to\infty]{\mathbb{P}} 0.$$

Indeed, the second part of Theorem 3 of [7] shows that these conditions hold for the reduced trees $(\{\xi(i), 1\leq i\leq k+1\}, d)$. As a consequence, the family $(\mathcal{R}(k), k\in\mathbb{N})$ can be represented as a continuous random tree $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T})$, and $(\delta(\xi(i),\xi(j)))_{i,j\geq 0}$ is the matrix of mutual distances between the points of an i.i.d. sample of $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T})$. This tree $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T})$ is called the cut-tree of $\mathcal{T}$. Note that $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T})$ depends on $\mathcal{T}$ and on the extra randomness of the Poisson process.

1.3 Fragmentation and cut-tree of the Brownian tree

We will also work on the Brownian tree $(\mathcal{T}^{br}, d^{br}, \rho^{br})$, which was defined by Aldous (see [7]) as the $\mathbb{R}$-tree coded by $(H_t)_{0\leq t\leq 1} = (2B_t)_{0\leq t\leq 1}$, where B denotes the standard Brownian excursion of length 1. This tree can be seen as the stable tree of index $\alpha=2$ (up to a scale factor, with the normalization we will use). In particular, we have a probability mass-measure μ^{br} on $\mathcal{T}^{br}$, defined as the image of the Lebesgue measure on $[0,1]$ under the canonical projection. We also define a length-measure l on $\mathcal{T}^{br}$, which is the sigma-finite measure such that, for all $u,v\in\mathcal{T}^{br}$, $l([\![u,v]\!]) = d^{br}(u,v)$.

The fragmentation of the Brownian tree we consider is the same as in [19]: conditionally on $\mathcal{T}^{br}$, we let $(t_i,b_i)_{i\in I}$ be the family of the atoms of a Poisson point process with intensity $dt\otimes l(dv)$ on $\mathbb{R}_+\times\mathcal{T}^{br}$. As for the stable tree, we let $\mathcal{T}_x^{br}(t)$ be the connected component of $\mathcal{T}^{br}\setminus\{b_i : t_i\leq t\}$, and $\mu_x^{br}(t) = \mu^{br}(\mathcal{T}_x^{br}(t))$, for every $x\in\mathcal{T}^{br}$. Adding a distinguished point 0 to $\mathcal{T}^{br}$, we define a function δ^{br} on $(\mathcal{T}^{br}\sqcup\{0\})^2$ such that for all $x,y\in\mathcal{T}^{br}$,

$$\delta^{br}(0,0)=0, \qquad \delta^{br}(0,x)=\delta^{br}(x,0)=\int_0^\infty \mu_x^{br}(t)dt,$$

$$\delta^{br}(x,y)=\int_{t^{br}(x,y)}^\infty (\mu_x^{br}(t)+\mu_y^{br}(t))dt,$$

where $t^{br}(x,y) := \inf\left\{t\in\mathbb{R}_+ : \mathcal{T}_x^{br}(t)\neq\mathcal{T}_y^{br}(t)\right\}$ is a.s. finite. As shown in [19], we can define a new tree $\mathrm{Cut}(\mathcal{T}^{br})$ for which the matrix of mutual distances between the points of an i.i.d. sample of $\mathrm{Cut}(\mathcal{T}^{br})$ is $(\delta(\xi(i),\xi(j)))_{i,j\geq 0}$, where $\xi(0)=0$ and $(\xi(i))_{i\in\mathbb{N}}$ is an i.i.d. sequence distributed according to μ^{br}, conditionally on $\mathcal{T}^{br}$. Moreover, $\mathrm{Cut}(\mathcal{T}^{br})$ has the same law as $\mathcal{T}^{br}$.

1.4 Main results

As stated in the Introduction, we mainly work in the setting of Galton–Watson trees with critical offspring distribution ν, where ν is a probability distribution belonging to the domain of attraction of a stable law of index $\alpha\in(1,2)$. We shall also assume that ν is aperiodic. Finally, for a technical reason, we will need the additional hypothesis

$$\sup_{r\geq 1}\left(\frac{r\mathbb{P}(\hat{Z}=r)}{\mathbb{P}(\hat{Z}>r)}\right)<\infty, \tag{2.1}$$

where $\hat{Z}$ is a random variable such that $\mathbb{P}(\hat{Z} = r) = r\nu(\{r\})$. For example, this is the case if $\nu(\{r\})$ is equivalent to $c/r^{\alpha+1}$ as $n \to \infty$, for a constant $c \in (0, \infty)$. In all our work, we shall implicitly work for values of n such that, for a Galton–Watson tree T with offspring distribution ν, $\mathbb{P}(|E(T)| = n) \neq 0$. We let $\mathcal{T}_n$ be a ν-Galton–Watson tree, conditioned to have exactly n edges. We let δ_n denote the graph-distance on $\{0, 1, \ldots, n\}$ induced by $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T}_n)$. We will use the notation ρ_n for the root of $\mathcal{T}_n$, and μ_n for the uniform distribution on $E(\mathcal{T}_n)$ (by slight abuse, μ_n will also sometimes be used for the uniform distribution on $\{1, \ldots, n\}$).

Our main goal is to study the asymptotic behavior of $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T}_n)$ as $n \to \infty$. To this end, it will be convenient to see trees as pointed metric measure spaces, and work with the Gromov–Prokhorov topology on the set of (equivalence classes of) such spaces. Let us recall a few definitions and facts on these objects (see for example [36] for details).

A pointed metric measure space is a quadruple (X, D, m, x), where m is a Borel probability measure on the metric space (X, D), and x is a point of X. These objects are considered up to a natural notion of isometry-equivalence. One says that a sequence (X_n, D_n, m_n, x_n) of pointed measure metric spaces converges in the Gromov–Prokhorov sense to $(X_\infty, D_\infty, m_\infty, x_\infty)$ if and only if the following holds: For $n \in \mathbb{N} \cup \{\infty\}$, set $\xi_n(0) = x_n$ and let $\xi_n(1), \xi_n(2), \ldots$ be a sequence of i.i.d. random variables with law m_n, then the vector $(D_n(\xi_n(i), \xi_n(j)) : 0 \leq i, j \leq k)$ converges in distribution to $(D_\infty(\xi_\infty(i); \xi_\infty(j)) : 0 \leq i, j \leq k)$ for every $k \geq 1$. The space $\mathbb{M}$ of (isometry-equivalence classes of) pointed measure metric spaces, endowed with the Gromov–Prokhorov topology, is a Polish space.

In this setting, the stable tree $\mathcal{T}$ with index α can be seen as a scaling limit of the Galton–Watson trees $\mathcal{T}_n, n \in \mathbb{N}$. More precisely, we endow the discrete trees $\mathcal{T}_n$ with the associated graph-distance d_n and the uniform distribution m_n on $V(\mathcal{T}_n) \setminus \{\rho_n\}$. Note that m_n is uniform on $\{v_1(\mathcal{T}_n), \ldots, v_n(\mathcal{T}_n)\}$; by slight abuse, it will sometimes be identified with the uniform distribution on $\{1, \ldots, n\}$. For any pointed metric measure space $\mathbf{X} = (X, D, m, x)$ and any $a \in (0, \infty)$, we let $a\mathbf{X} = (X, aD, m, x)$. With this formalism, there exists a sequence $(a_n)_{n \in \mathbb{N}}$ such that

$$\frac{a_n}{n} \mathcal{T}_n \xrightarrow{(d)} \mathcal{T}, \tag{2.2}$$

in the sense of the Gromov–Prokhorov topology, and $a_n = n^{1/\alpha} f(n)$ for a slowly-varying function f. This is a consequence of the convergence of the contour functions associated with the trees $\mathcal{T}_n$, shown in [31, Theorem 3.1]. We will give a slightly more precise version of this result in Section 2.2.2.

We can now state our main result:

Theorem 1.3. *Let $(a_n)_{n \in \mathbb{N}}$ be a sequence such that (2.2) holds. Then we have the following joint convergence in distribution:*

$$\left(\frac{a_n}{n} \mathcal{T}_n, \frac{a_n}{n} \mathrm{Cut}_{\mathrm{v}}(\mathcal{T}_n) \right) \xrightarrow[n \to \infty]{} (\mathcal{T}, \mathrm{Cut}_{\mathrm{v}}(\mathcal{T})) \,,$$

where $\mathbb{M}$ is endowed with the Gromov–Prokhorov topology and $\mathbb{M} \times \mathbb{M}$ has the associated product topology. Furthermore, the cut-tree $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T})$ has the same distribution as $\mathcal{T}$.

Note that this generalizes Proposition 1.4 of [1], which gave the scaling limit of the number of cuts needed to isolate the root in a stable Galton–Watson tree.

In the following sections, we fix the sequence (a_n). For some of the preliminary results, we will use a particular choice of this sequence, detailed in Section 2.2.1. Nevertheless, it is easy to check that the theorem holds for any equivalent sequence.

To complete this result, we will study the limit of the cut-tree obtained for the vertex-fragmentation, in the case where the offspring distribution ν has finite variance (still assuming that ν is critical and aperiodic). More precisely, we will show the following:

Theorem 1.4. *If the offspring distribution ν has finite variance σ^2, then we have the joint convergence in distribution*

$$\left(\frac{\sigma}{\sqrt{n}}\mathcal{T}_n, \frac{1}{\sqrt{n}}\left(\sigma+\frac{1}{\sigma}\right)\mathrm{Cut}_{\mathrm{v}}(\mathcal{T}_n)\right) \xrightarrow[n\to\infty]{} \left(\mathcal{T}^{br}, \mathrm{Cut}(\mathcal{T}^{br})\right)$$

in $\mathbb{M}\times\mathbb{M}$.

Let us explain informally why we get a factor $\sigma+1/\sigma$, instead of the $1/\sigma$ we had in the case of the edge-fragmentation. In the vertex-fragmentation, the average number of deleted edges at each step is roughly $\sum_k k\nu(k)\times k = \sigma^2+1$. Thus, the edge-deletions happen σ^2+1 times faster than for the edge-fragmentation. As a consequence, $(1/\sqrt{n})\cdot \mathrm{Cut}_{\mathrm{v}}(\mathcal{T}_n)$ behaves approximatively like $(1/(\sigma^2+1)\sqrt{n})\cdot\mathrm{Cut}(\mathcal{T}_n)$, i.e. $(\sigma+1/\sigma)^{-1}(1/\sigma\sqrt{n})\cdot\mathrm{Cut}(\mathcal{T}_n)$.

Also note that we would need additional hypotheses to extend this result to the more general case of an offspring distribution belonging to the domain of attraction of a Gaussian distribution. Indeed, as will be seen in the Section 4, the proof of this result relies on the convergence of the coefficients n/a_n^2: if ν has finite variance, we may and will take $a_n=\sigma\sqrt{n}$, but in the general case, this convergence is not granted.

For both of these theorems, it is known that the first component converges in the stronger sense of the Gromov–Hausdorff–Prokhorov topology. However, as in the case studied by Bertoin and Miermont, the question of whether the joint convergences hold in this sense remains open.

In the following sections, we will first work on the proof of Theorem 1.3: preliminary results will be given in Section 2, and the proof will be completed in Section 3. The global structure of this proof is close to that of [19], although the technical arguments differ, especially in Section 2. Section 4 will be devoted to the study of the finite variance case.

2 Preliminary results

2.1 Modified distance on $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T}_n)$

We begin by introducing a new distance δ'_n on $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T}_n)$, defined in a similar way as the distance δ for a continuous tree. We show that this distance is "close" enough to $(a_n/n)\cdot\delta_n$, which will enable us to work on the modified cut-tree $\mathrm{Cut}'_{\mathrm{v}}(\mathcal{T}_n) := (\mathrm{Cut}_{\mathrm{v}}(\mathcal{T}_n), \delta'_n)$.

Recall the fragmentation of $\mathcal{T}_n$ introduced in Section 1.1. We now turn this process into a continuous-time fragmentation, by saying that each vertex $v\in V(T)$ is marked independently, with rate $\deg v/a_n$. Equivalently, this can be seen as marking each edge of T independently with rate $1/a_n$, and deleting all the edges e such that $e^- = e_i^-$ as soon as e_i is marked. Thus, we obtain a forest $\overline{\mathcal{T}}_n(t)$ at time t. For every $i\in\{1,\ldots,n\}$, we let $\mathcal{T}_{n,i}(t)$ denote the component of $\overline{\mathcal{T}}_n(t)$ containing the edge e_i, with the convention $\mathcal{T}_{n,i}(t)=\emptyset$ if $e_i\notin\overline{\mathcal{T}}_n(t)$, and $\mu_{n,i}(t)=\mu_n(\mathcal{T}_{n,i}(t))$. Note that $n\mu_{n,i}(t)$ is the number of edges in $\mathcal{T}_{n,i}(t)$. For all $i,j\in\{1,\ldots,n\}$, we now define

$$\delta'_n(0,0)=0, \quad \delta'_n(0,i)=\delta'_n(i,0)=\int_0^\infty \mu_{n,i}(t)dt$$
$$\delta'_n(i,j)=\int_{t_n(i,j)}^\infty (\mu_{n,i}(t)+\mu_{n,j}(t))\,dt,$$

where $t_n(i,j)$ denotes the first time when the components $\mathcal{T}_{n,i}(t)$ and $\mathcal{T}_{n,j}(t)$ become disjoint.

Lemma 2.1. *For all $i,j\in\{1,\ldots,n\}$, we have*

$$\mathbb{E}\left[\left|\frac{a_n}{n}\delta_n(0,i)-\delta'_n(0,i)\right|^2\right]=\frac{a_n}{n}\mathbb{E}\left[\delta'_n(0,i)\right]$$

and

$$\mathbb{E}\left[\left|\frac{a_n}{n}\delta_n(i,j)-\delta'_n(i,j)\right|^2\right]\le\frac{a_n}{n}\mathbb{E}\left[\delta'_n(0,i)+\delta'_n(0,j)\right].$$

Proof. We work conditionally on $\mathcal{T}_n$. Fix $i \in \{1, \dots, n\}$. For all $t \in \mathbb{R}_+$, we let $N_i(t)$ be the number of cuts happening in the component containing e_i up to time t. Since each edge of $\mathcal{T}_n$ is marked independently with rate $1/a_n$, the process $(M_i(t))_{t \geq 0}$, where

$$M_i(t) := \frac{a_n}{n} N_i(t) - \int_0^t \mu_i(s) ds,$$

is a purely discontinuous martingale. Its predictable quadratic variation can be written as

$$\langle M_i \rangle_t = \frac{a_n}{n} \int_0^t \mu_i(s) ds.$$

As a consequence, we have $\mathbb{E}[|M_i(\infty)|^2] = \mathbb{E}\left[\langle M_i \rangle_\infty\right]$. Since

$$\lim_{t \to \infty} N_i(t) = \delta_n(0,i) \quad \text{and} \quad \lim_{t \to \infty} \int_0^t \mu_i(s) ds = \delta'_n(0,i),$$

we get

$$\mathbb{E}\left[\left|\frac{a_n}{n}\delta_n(0,i) - \delta'_n(0,i)\right|^2\right] = \frac{a_n}{n}\mathbb{E}\left[\delta'_n(0,i)\right].$$

For the second part, we use similar arguments. We fix $i \neq j \in \{1, \dots, n\}$, and we write t_{ij} instead of $t_n(i,j)$. For all $t \geq 0$, let $\mathcal{F}_t$ denote the σ-algebra generated by $\mathcal{T}_n$ and the atoms $\{(t_r, e_{i_r}) : t_r \leq t\}$ of the Poisson point process of marks on the edges introduced in Section 1.1. Conditionally on $\mathcal{F}_{t_{ij}}$,

$$M_{ij}(t) := M_i(t_{ij} + t) - M_i(t_{ij}) + M_j(t_{ij} + t) - M_j(t_{ij})$$

defines a purely discontinuous martingale such that

$$\begin{aligned}
\lim_{t \to \infty} M_{ij}(t) &= \frac{a_n}{n}(\delta_n(b_{ij}, i) + \delta_n(b_{ij}, j)) - \int_{t_{ij}}^{\infty} \mu_i(s) ds - \int_{t_{ij}}^{\infty} \mu_j(s) ds \\
&= \frac{a_n}{n}\delta_n(i,j) - \delta'_n(i,j),
\end{aligned}$$

where b_{ij} denotes the most recent common ancestor of the leaves i and j in $\mathrm{Cut_v}(\mathcal{T}_n)$. Besides, since the edges of $\mathcal{T}_{n,i}$ and $\mathcal{T}_{n,j}$ are marked independently after time t_{ij}, the predictable quadratic variation of M_{ij} is

$$\langle M_{ij} \rangle_t = \frac{a_n}{n}\mathbb{E}\left[\int_{t_{ij}}^{t_{ij}+t} (\mu_i(s) + \mu_j(s))\, ds\right].$$

Since $\delta'_n(i,j) = \delta'_n(0,i) + \delta'_n(0,j) - 2\delta'_n(0, b_{ij})$, this yields

$$\mathbb{E}\left[\left|\frac{a_n}{n}\delta_n(i,j) - \delta'_n(i,j)\right|^2\right] \leq \frac{a_n}{n}\mathbb{E}\left[\delta'_n(0,i) + \delta'_n(0,j)\right].$$

□

2.2 A first joint convergence

In this section, we first state precisely the convergence theorems we will rely on to prove the following lemmas. To this end, we work in the setting of sums of i.i.d. random variable $S_n = Z_1 + \ldots + Z_n$, where the laws of the Z_i are in the domain of attraction of a stable law. Under additional hypotheses, Theorem 2.2 below gives a choice of scaling constants a_n for which S_n/a_n converges in law to a stable variable, and a formulation of Gnedenko's local limit theorem in this setting. Next, we will recall a result of Duquesne which shows, in particular, the convergence (2.2). The version we will use is a joint convergence of three functions encoding the trees $\mathcal{T}_n$ and $\mathcal{T}$. These results will allow us to prove a first joint convergence for the fragmented trees in Proposition 2.5.

2.2.1 Local limit theorem

We say that a measure π on $\mathbb{Z}$ is lattice if there exists integers $b \in \mathbb{Z}$, $d \geq 2$ such that $\operatorname{supp}(\pi) \subset b + d\mathbb{Z}$. We know from our hypotheses that ν is critical, aperiodic, and $\nu(\{0\}) > 0$, and these three conditions imply that ν is nonlattice.

For any $\beta \in (1,2)$, we let $X^{(\beta)}$ be a stable spectrally positive Lévy process with parameter β, and $p_t^{(\beta)}(x)$ the density of the law of $X_t^{(\beta)}$. Similarly, for $\beta \in (0,1)$, we let $X^{(\beta)}$ be a stable subordinator with parameter β, and $q_t^{(\beta)}(x)$ be the density of the law of $X_t^{(\beta)}$. We fix the normalization of these processes by setting, for all $\lambda \geq 0$,

$$\mathbb{E}\left[e^{-\lambda X_t^{(\beta)}}\right] = e^{t\lambda^\beta} \qquad \text{if } \beta \in (1,2)$$
$$\mathbb{E}\left[e^{-\lambda X_t^{(\beta)}}\right] = e^{-t\lambda^\beta} \qquad \text{if } \beta \in (0,1).$$

We also introduce the set R_ρ of regularly varying functions with index ρ.

Theorem 2.2. *Let $(Z_i, i \in \mathbb{N})$ be an i.i.d. sequence of random variables in $\mathbb{N} \cup \{-1, 0\}$. We denote by Z a random variable having the same law as the Z_i. Suppose that the law of Z belongs to the domain of attraction of a stable law of index $\beta \in (0,2) \setminus \{1\}$, and is nonlattice. If $\beta \in (1,2)$, we also suppose that Z is centered. We introduce*

$$S_n = \sum_{i=1}^{n} Z_i, \quad n \geq 0.$$

Then there exists an increasing function $A \in R_\beta$ and a constant c such that

(i) It holds that

$$\mathbb{P}(Z > r) \sim \frac{c}{A(r)} \quad \text{as } r \to \infty. \tag{2.3}$$

(ii) Letting a be the inverse function of A, and $a_n = a(n)$ for all $n \in \mathbb{N}$, we have

$$\lim_{n\to\infty} \sup_{k \in \mathbb{N}} \left| a_n \mathbb{P}(S_n = k) - p_1^{(\beta)}\left(\frac{k}{a_n}\right)\right| = 0. \tag{2.4}$$

Proof. Theorem 8.3.1 of [20] shows that, since $Z \geq -1$ a.s., the law of Z belongs to the domain of attraction of a stable law of index β if and only if $\mathbb{P}(Z > r) \in R_{-\beta}$. Using Theorem 1.5.3 of [20], we can take a monotone equivalent of $\mathbb{P}(Z > r)$, hence the existence of A such that (2.3) holds with a constant c which will be chosen hereafter.

The remarks following Theorem 8.3.1 in [20] give a characterization of the a_n such that S_n/a_n converges in law to a stable variable of index β. In particular, it is enough to take a_n such that $n/A(a_n)$ converges, so $a = A^{-1}$ is a suitable choice. We now choose the constant c such that S_n/a_n converges to $X_1^{(\beta)}$. The second point of the theorem is given by Gnedenko's local limit theorem (see, e.g., Theorem 4.2.1 of [39]). □

2.2.2 Coding the trees $\mathcal{T}_n$ and $\mathcal{T}$

We now recall three classical ways of coding a tree $T \in \mathbb{T}$, namely the associated contour function, height function and Lukasiewicz path. Detailed descriptions and properties of these objects can be found, for example, in [31].

To define the contour function $C^{[n]}$ of $\mathcal{T}_n$, we see $\mathcal{T}_n$ as the embedded tree in the oriented half-plane, with each edge having length 1. We consider a particle that visits continuously all edges at unit speed, from the left to the right, starting from the root. Then, for every $t \in [0, 2n]$, we let $C_t^{[n]}$ be the *height* of the particle at time t, that is, its distance to the root. The height function is defined by letting $H_j^{[n]}$ be the height of the vertex v_j. Lastly, for all $i \in \{0, \ldots, n\}$,

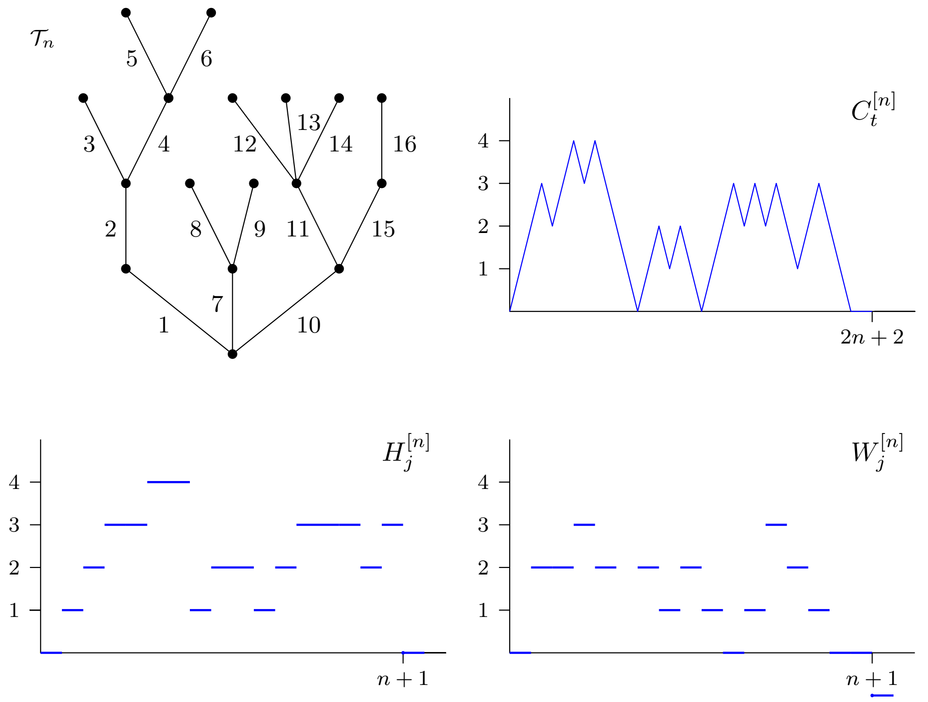

Figure 2.2 – The contour function $(C_t^{[n]}, 0 \le t \le 2n+2)$, height function $(H_j^{[n]}, j = 0, \ldots, n+1)$ and Lukasiewicz path $(W_j^{[n]}, j = 0, \ldots, n+1)$ coding a realization of $\mathcal{T}_n$.

we let $Z_{i+1}^{[n]}$ be the number of offspring of the vertex v_i. Then the Lukasiewicz path of $\mathcal{T}_n$ is defined by

$$W_j^{[n]} = \sum_{i=1}^{j} Z_i^{[n]} - j, \quad j = 0, \ldots, n+1.$$

With this definition, we have $\deg(v_j, \mathcal{T}_n) = W_{j+1}^{[n]} - W_j^{[n]} + 1$. We extend $C^{[n]}$ and $H^{[n]}$ by setting $C_t^{[n]} = 0$ for all $t \in [2n, 2n+2]$ and $H_{n+1}^{[n]} = 0$ (this will allow us to keep similar scaling factors for the rescaled functions we introduce in Theorem 2.3). Figure 2.2 gives the contour function, height function and Lukasiewicz path associated to the tree we used in Figure 2.1.

We also use a random walk $(W_j)_{j \ge 0}$ with jump distribution $\nu(k+1)$:

$$W_j = \sum_{i=1}^{j} Z_i - j, \quad j \ge 0,$$

where $(Z_i)_{i \in \mathbb{N}}$ are i.i.d. variables having law ν. Note that $(W_j^{[n]}, j = 0, \ldots, n+1)$ has the same law as $(W_j, j = 0, \ldots, n+1)$ conditionally on $W_{n+1} = -1$ and $W_j \ge 0$ for all $j \le n$. In other terms, $(W_n)_{n \ge 0}$ has the same law as the Lukasiewicz path associated with a sequence of Galton–Watson trees with offspring distribution ν. From now on, we let A and a be functions given by Theorem 2.2 for the sequence of i.i.d. variables $(Z_i - 1)_{i \in \mathbb{N}}$. Thus, we have the convergence

$$\frac{1}{a_n} W_n \xrightarrow[n \to \infty]{(d)} X_1^{(\alpha)}. \tag{2.5}$$

Finally, let $(X_t)_{0 \le t \le 1}$ be the excursion of length 1 of the Lévy process $X^{(\alpha)}$, and $(H_t)_{0 \le t \le 1}$ be the excursion of length 1 of the process $H^{(\alpha)}$ defined in Section 1.2. We will use the following adaptation of the results shown by Duquesne in [31]:

Theorem 2.3 (Duquesne). *Consider the rescaled functions $C^{(n)}$, $H^{(n)}$ and $X^{(n)}$, defined by*

$$C_t^{(n)} = \frac{a_n}{n} C^{[n]}_{(2n+2)t}, \qquad H_t^{(n)} = \frac{a_n}{n} H^{[n]}_{\lfloor (n+1)t \rfloor}, \qquad X_t^{(n)} = \frac{1}{a_n} W^{[n]}_{\lfloor (n+1)t \rfloor}$$

for all $t \in [0,1]$. If ν is aperiodic and hypothesis (2.5) *holds, then we have the joint convergence*

$$\left(C_t^{(n)}, H_t^{(n)}, X_t^{(n)}\right)_{0 \le t \le 1} \xrightarrow[n\to\infty]{(d)} (H_t, H_t, X_t)_{0\le t\le 1}\,.$$

Proposition 4.3 of [31] shows the convergence of the corresponding bridges (with a change of index which comes from the fact that we are working on trees conditioned to have n edges instead of n vertices). Using the continuity of the Vervaat transform as in the proof of [31, Theorem 3.1] then gives the result.

The fact that these convergences hold jointly will be used in the proof of Lemma 2.4 below. Apart from this, we will mainly use the convergence of the rescaled Lukasiewicz paths $X^{(n)}$, because of the following link between the rates of our fragmentation and the jumps of $X^{(n)}$. Recall from Section 1.2 that $p : [0,1] \to \mathcal{T}$ denotes the canonical projection from $[0,1]$ onto $\mathcal{T}$. Now, the set of the branching points of $\mathcal{T}$ is $\{p(t) : t \in [0,1] \text{ s.t. } \Delta X_t > 0\}$, and the associated local times are $L(p(t)) = \Delta X_t$ (see [33, proof of Theorem 4.7] and [50, Proposition 2]). Similarly, we introduce the projection p_n from $K_n := \{1/(n+1), \dots, 1\}$ onto $V(\mathcal{T}_n)$, such that $p_n(j/(n+1))$ is the vertex v_{j-1} of $\mathcal{T}_n$. Thus, for all $t \in K_n$, we have

$$\Delta X_t^{(n)} = \frac{1}{a_n}(\deg(p_n(t), \mathcal{T}_n) - 1). \tag{2.6}$$

We conclude this part by showing another result of joint convergence, for the Lukasiewicz paths of two symmetric sequences of trees. For all $n \in \mathbb{N}$, we introduce the symmetrized tree $\tilde{\mathcal{T}}_n$, obtained by reversing the order of the children of each vertex of $\mathcal{T}_n$. We let $\tilde{W}^{[n]}$ denote the Lukasiewicz path of $\tilde{\mathcal{T}}_n$. (We would obtain the same process by visiting the vertices of $\mathcal{T}_n$ "from right to left" in the depth-first search.) Finally, we define the rescaled process $\tilde{X}^{(n)}$ by

$$\tilde{X}_t^{(n)} = \frac{1}{a_n} \tilde{W}^{[n]}_{\lfloor (n+1)t \rfloor} \qquad \forall t \in [0,1].$$

Lemma 2.4. *There exists a process $(\tilde{X}_t)_{0\le t\le 1}$ such that there is the joint convergence*

$$(X^{(n)}, \tilde{X}^{(n)}) \xrightarrow[n\to\infty]{(d)} (X, \tilde{X}). \tag{2.7}$$

Moreover :
— *the processes $\tilde{X}$ and X have the same law.*
— *for every jump-time t of X,*

$$\Delta \tilde{X}_{1-t-l(t)} = \Delta X_t \qquad a.s.,$$

where $l(t) = \inf\{s > t : X_s = X_{t-}\} - t$.

Proof. Since $\mathcal{T}_n$ and $\tilde{\mathcal{T}}_n$ have the same law, $\tilde{X}^{(n)}$ converges in distribution to an excursion of the Lévy process $X^{(\alpha)}$ in the Skorokhod space $\mathbb{D}$. Thus the sequence of the laws of the processes $(X^{(n)}, \tilde{X}^{(n)})$ is tight in $\mathbb{D}\times\mathbb{D}$. Up to extraction, we can assume that $(X^{(n)}, \tilde{X}^{(n)})$ converges in distribution to a couple of processes $(X, \tilde{X})$.

For all $n \in \mathbb{N}$, $j \in \{0, \dots, n\}$, a simple computation shows that the vertex $v_j(\mathcal{T}_n)$ corresponds to $v_{\tilde{j}}(\tilde{\mathcal{T}}_n)$, where

$$\tilde{j} = n - j + H_j^{[n]} - D_j^{[n]},$$

and $D_j^{[n]}$ is the number of strict descendants of $v_j(\mathcal{T}_n)$. Note that $D_j^{[n]}$ is the largest integer such that $W_i^{[n]} \geq W_j^{[n]}$ for all $i \in [j, j + D_j^{[n]}]$. Then (2.6) shows that we have

$$\Delta \tilde{X}^{(n)}_{(n-j+H_j^{[n]}-D_j^{[n]}+1)/(n+1)} = \Delta X^{(n)}_{(j+1)/(n+1)}. \tag{2.8}$$

For all $n \in \mathbb{N} \cup \{\infty\}$, we let $(s_i^{(n)})_{i\in\mathbb{N}}$ be the sequence of the times where $X^{(n)}$ has a positive jump, ranked in such a way that the sequence of the jumps $(\Delta X^{(n)}_{s_i^{(n)}})_{i\in\mathbb{N}}$ is nonincreasing. We define the $(\tilde{s}_i^{(n)})_{i\in\mathbb{N}}$ in a similar way for the $\tilde{X}^{(n)}$, $n \in \mathbb{N} \cup \{\infty\}$. Fix $i \in \mathbb{N}$. Then (2.8) can be translated into

$$\tilde{s}_i^{(n)} = 1 - s_i^{(n)} + \frac{1}{n+1}\left(1 + H^{[n]}_{(n+1)s_i^{(n)}-1} - D^{[n]}_{(n+1)s_i^{(n)}-1}\right). \tag{2.9}$$

Using the Skorokhod representation theorem, we now work under the hypothesis

$$\left(H_t^{(n)}, X_t^{(n)}\right)_{0\leq t\leq 1} \xrightarrow[n\to\infty]{} (H_t, X_t)_{0\leq t\leq 1} \qquad \text{a.s.}$$

Then the following convergences hold a.s., for all $i \geq 1$:

$$\begin{aligned} s_i^{(n)} &\xrightarrow[n\to\infty]{} s_i \\ \Delta X^{(n)}_{s_i^{(n)}} &\xrightarrow[n\to\infty]{} \Delta X_{s_i} \\ \frac{1}{n+1} H^{[n]}_{(n+1)s_i^{(n)}-1} &\xrightarrow[n\to\infty]{} 0 \\ \frac{1}{n+1} D^{[n]}_{(n+1)s_i^{(n)}-1} &\xrightarrow[n\to\infty]{} l(s_i). \end{aligned}$$

The first two convergences hold because the ΔX_{s_i} are distinct, and the last one uses the fact that a.s.

$$\inf_{0\leq u\leq\varepsilon} X_{s_i+l(s_i)+u} < X_{(s_i)^-} \qquad \forall\varepsilon > 0.$$

As a consequence, $\tilde{s}_i^{(n)}$ converges a.s. to $1 - s_i - l(s_i)$. Thus $\tilde{s}_i = 1 - s_i - l(s_i)$ a.s., and $\Delta\tilde{X}_{\tilde{s}_i} = \Delta X_{s_i}$ a.s. (Since the discontinuity points are countable, this holds jointly for all i).

The Lévy–Itô representation theorem shows that $\tilde{X}$ can be written as a measurable function of $(\tilde{s}_i, \Delta\tilde{X}_{\tilde{s}_i})_{i\in\mathbb{N}}$. This identifies uniquely the law of $(X, \tilde{X})$, hence (2.7). □

2.2.3 Joint convergence of the subtree sizes

Recall from Section 1.2 that $(\xi(i), i \in \mathbb{N})$ is a sequence of i.i.d. variables in $\mathcal{T}$, with distribution the mass-measure μ, and $\xi(0) = 0$. For all $n \in \mathbb{N}$, we introduce independent sequences $(\xi_n(i), i \in \mathbb{N})$ of i.i.d. uniform integers in $\{1, \dots, n\}$, and set $\xi_n(0) = 0$. Recalling the notation of Section 2.1, we let $\tau_n(i,j) = t_n(\xi_n(i), \xi_n(j))$ be the first time when the components $\mathcal{T}_{n,\xi_n(i)}(t)$ and $\mathcal{T}_{n,\xi_n(j)}(t)$ become disjoint. Similarly $\tau(i,j)$ will denote the first time when the components containing $\xi(i)$ and $\xi(j)$ become disjoint in the fragmentation of $\mathcal{T}$. Our goal is to prove the following result.

Proposition 2.5. *As $n \to \infty$, we have the following weak convergences*

$$\begin{gathered} \frac{a_n}{n}\mathcal{T}_n \xrightarrow{(d)} \mathcal{T} \\ (\tau_n(i,j))_{i,j\in\mathbb{N}} \xrightarrow{(d)} (\tau(i,j))_{i,j\in\mathbb{N}} \\ \left(\mu_{n,\xi_n(i)}(t)\right)_{i\in\mathbb{N},t\geq 0} \xrightarrow{(d)} \left(\mu_{\xi(i)}(t)\right)_{i\in\mathbb{N},t\geq 0}, \end{gathered}$$

where the three hold jointly.

For the proof of this proposition, it will be convenient to identify the $\xi_n(i)$ with vertices of $\mathcal{T}_n$ instead of edges. As noted in [19, proof of Lemma 2], this makes no difference for the result we seek.

We let

$$t_i^{(n)} = \frac{\xi_n(i)+1}{n+1},$$

so that $p_n(t_i^{(n)}) = v_{\xi_n(i)}(\mathcal{T}_n)$. Furthermore, we may and will take $\xi(i) = p(t_i)$, with a sequence $(t_i, i \in \mathbb{N})$ of independent uniform variables in $[0,1]$. The sequence $(t_i^{(n)}, i \in \mathbb{N})$ converges in distribution to $(t_i, i \in \mathbb{N})$. Since these sequences are independent of the trees $\mathcal{T}_n$ and $\mathcal{T}$, the Skorokhod representation theorem allows us to assume

$$\begin{cases} \left(X^{(n)}, \tilde{X}^{(n)}\right) \xrightarrow[n\to\infty]{} \left(X, \tilde{X}\right) & \text{a.s.} \\ \left(t_i^{(n)}, i \in \mathbb{N}\right) \xrightarrow[n\to\infty]{} (t_i, i \in \mathbb{N}) & \text{a.s.} \end{cases} \tag{2.10}$$

We will sometimes write $X_t^{(\infty)}$ and $t_i^{(\infty)}$ for X_t and t_i, when it makes notation easier.

For any two vertices u, v of a discrete tree T, we introduce the notation

$$[\![u,v]\!]_V = [\![u,v]\!] \cap V(T) \qquad \text{and} \qquad]\!]u,v[\![_V = [\![u,v]\!]_V \setminus \{u,v\},$$

where $[\![u,v]\!]$ is the segment between u and v in T (seen as an $\mathbb{R}$-tree).

Definition 2.6. *Fix $T \in \mathbb{T}$. The shape of T is the discrete tree $S(T)$ such that*

$$\begin{aligned} V(S(T)) &= \{v \in V(T) : \deg v \neq 1\} \\ E(S(T)) &= \left\{\{u,v\} \in V(S(T))^2 : \forall w \in]\!]u,v[\![_V,\ \deg w = 1\right\}. \end{aligned}$$

Note that this definition can easily be extended to the case of an $\mathbb{R}$-tree (T,d) having a finite number of leaves, by using the "convention" $V(T) = \{v \in T : \deg v \neq 1\}$ in the previous definition.

For all $n, k \in \mathbb{N}$, we let $\mathcal{R}_n(k)$ denote the shape of the subtree of $\mathcal{T}_n$ spanned by the vertices $\xi_n(1), \ldots, \xi_n(k)$ and the root. Similarly, $\mathcal{R}_\infty(k)$ will denote the shape of the subtree of $\mathcal{T}$ spanned by $\xi(1), \ldots, \xi(k)$ and the root. For all $n \in \mathbb{N} \cup \{\infty\}$, we let $V_n(k)$ be the set of the vertices of $\mathcal{R}_n(k)$, and we identify the edges of $\mathcal{R}_n(k)$ with the corresponding segments in $\mathcal{T}_n$. In particular, for any edge $e = \{u,v\}$ of $\mathcal{R}_n(k)$, we write $w \in e$ if $w \in]\!]u,v[\![_V$. We let $L_n(v)$ denote the rate at which a vertex v is deleted in $\mathcal{T}_n$. Recall from Section 2.1 that $L_n(v) = \deg(v, \mathcal{T}_n)/a_n$.

Lemma 2.7. *Fix $k \in \mathbb{N}$. Under (2.10), $\mathcal{R}_n(k)$ is a.s. constant for all n large enough (say $n \geq N$). Identifying $V_n(k)$ with $V_\infty(k)$ for all $n \geq N$, we have*

$$(L_n(v), v \in V_n(k)) \xrightarrow[n\to\infty]{} (L(v), v \in V_\infty(k)) \qquad \textit{a.s.}$$

The above convergence can be written more rigorously by numbering the vertices of $\mathcal{R}_n(k)$ and $\mathcal{R}_\infty(k)$, and indexing on $i \in \{1, \ldots, |V_\infty(k)|\}$, but we keep this form to make the notation easier.

Proof. For all $n \in \mathbb{N} \cup \{\infty\}$, $s < t \in [0,1]$, we let

$$I_{s,t}^{(n)} = \inf_{s<u<t} X_u^{(n)},$$

and for all $i, j \in \mathbb{N}$,

$$t_{ij}^{(n)} = \sup\left\{s \in \left[0, t_i^{(n)} \wedge t_j^{(n)}\right] : I^{(n)}_{s,t_i^{(n)}} = I^{(n)}_{s,t_j^{(n)}}\right\}.$$

Note that $p_n(t_{ij}^{(n)})$ is the most recent common ancestor of the vertices $\xi_n(i)$ and $\xi_n(j)$ in $\mathcal{T}_n$. If for example $t_i^{(n)} < t_j^{(n)}$, we can rewrite $t_{ij}^{(n)}$ as

$$\sup\left\{s \in \left[0, t_i^{(n)}\right] : X_{s-}^{(n)} \leq I_{t_i^{(n)}, t_j^{(n)}}^{(n)}\right\}.$$

Besides, for $n = \infty$, we can replace the inequality in the broad sense by a strict inequality:

$$t_{ij} = \sup\left\{s \in [0, t_i] : X_{s-} < I_{t_i, t_j}\right\}.$$

With this notation, it is elementary to show that the following properties hold a.s. for all $i, j, i', j' \geq 0$:

(i) X is continuous at t_i, and $X_{t_i^{(n)}}^{(n)}$ converges to X_{t_i} as $n \to \infty$.

(ii) $t_{ij}^{(n)}$ converges to t_{ij} as $n \to \infty$.

(iii) $X_{t_{ij}^{(n)}}^{(n)}$ converges to $X_{t_{ij}}$ and $X_{(t_{ij}^{(n)})-}^{(n)}$ converges to $X_{(t_{ij})-}$ as $n \to \infty$.

(iv) If $t_{ij} = t_{i'j'}$, then $t_{ij}^{(n)} = t_{i'j'}^{(n)}$ for all n large enough.

We now fix $k \in \mathbb{N}$. We introduce the set

$$B_n(k) = \left\{t_i^{(n)} : i \in \{1, \ldots, k\}\right\} \cup \left\{t_{ij}^{(n)} : i, j \in \{1, \ldots, k\}\right\} \cup \{0\}$$

of the times coding the vertices of $\mathcal{R}_n(k)$. We let $N_n(k)$ be the number of elements of $B_n(k)$, and $b_i^{(n,k)}$ be the ith element of $B_n(k)$. Properties (i)-(iv) can be translated into the a.s. properties:

(i)' For n large enough, $N_n(k)$ is constant.

(ii)' For all $i \in \{1, \ldots, N_\infty(k)\}$,

$$\begin{aligned} b_i^{(n,k)} &\xrightarrow[n\to\infty]{} b_i^{(\infty,k)} \\ X_{b_i^{(n,k)}}^{(n)} &\xrightarrow[n\to\infty]{} X_{b_i^{(\infty,k)}} \\ X_{(b_i^{(n,k)})-}^{(n)} &\xrightarrow[n\to\infty]{} X_{(b_i^{(\infty,k)})-}. \end{aligned}$$

Moreover, $\mathcal{R}_n(k)$ and the $L_n(v)$, $v \in V_n(k)$, can be recovered in a simple way using $B_n(k)$ and the $X_b^{(n)}$, $b \in B_n(k)$:

— Construct a graph with vertices labeled by $B_n(k)$, the root having label 0.

— For every $b \in B_n(k) \setminus \{0\}$, let b' denote the largest $b'' < b$ such that $b'' \in B_n(k)$ and $X_{b''}^{(n)} \leq X_b^{(n)}$, then draw an edge between the vertices labeled b and b'.

— For each vertex v labeled by $b \in B_n(k)$, let $L_n(v) = \Delta X_b^{(n)} + 1/a_n$.

This entails the lemma. □

This first lemma allows us to control the rate at which fragmentations happen at the vertices of $\mathcal{R}_n(k)$. We now need another quantity for the fragmentations happening "on the branches" of $\mathcal{R}_n(k)$, that is, at vertices $v \in V(\mathcal{T}_n) \setminus V_n(k)$. For every $n \in \mathbb{N} \cup \{\infty\}$, we let

$$\sigma_n(t) = \sum_{\substack{0<s<t \\ X_{s-}^{(n)} < I_{st}^{(n)}}} \Delta X_s^{(n)} \qquad \forall t \in [0, 1].$$

If $n \in \mathbb{N}$, the quantity $a_n \sigma_n(t)$ is the sum of the quantities $\deg v - 1$ over all strict ancestors $v \neq \rho_n$ of $p_n(t)$ in $\mathcal{T}_n$. Similarly, $\sigma(t)$ is the (infinite) sum of the $L(v)$ for all branching points v of $\mathcal{T}$ that are on the path $[\![p(t), \rho]\!]$.

Lemma 2.8. *With the preceding notation, in the setting of* (2.10)*, for all* $i \in 1, \dots, N(k)$*, we have the convergence*

$$\sigma_n(b_i^{(n,k)}) \xrightarrow[n\to\infty]{} \sigma_\infty(b_i^{(\infty,k)}) \quad a.s.$$

Proof. We fix $i \in \mathbb{N}$, and let $b_n = b_i^{(n,k)}$ to simplify the notation. For all $n \in \mathbb{N} \cup \{\infty\}$, we write $\sigma_n(t) = \sigma_n^-(t) + \sigma_n^+(t)$, where

$$\sigma_n^+(t) = \sum_{\substack{0<s<t \\ X_{s-}^{(n)} < I_{st}^{(n)}}} \left(X_s^{(n)} - I_{s,t}^{(n)} \right)$$
$$\sigma_n^-(t) = \sum_{\substack{0<s<t \\ X_{s-}^{(n)} < I_{st}^{(n)}}} \left(I_{s,t}^{(n)} - X_{s-}^{(n)} \right).$$

For any s, t such that $0 < s < t$ and $X_{s-}^{(n)} < I_{st}^{(n)}$, the term $a_n(X_s^{(n)} - I_{s,t}^{(n)})$ corresponds to the number of children of $p_n(s)$ that are visited before $p_n(t)$ in the depth-first search, and $a_n(I_{s,t}^{(n)} - X_{s-}^{(n)})$ is the number of children of $p_n(s)$ that are visited after $p_n(t)$. Writing the same decomposition $\tilde{\sigma}_n(t) = \tilde{\sigma}_n^-(t) + \tilde{\sigma}_n^+(t)$ for the trees $\tilde{\mathcal{T}}_n$, and recalling (2.9), we thus get

$$\sigma_n^+(b_n) = \tilde{\sigma}_n^-(\tilde{b}_n),$$

where

$$\tilde{b}_n = 1 - b_n + \frac{1}{n+1} \left(1 + H^{[n]}_{(n+1)b_n - 1} - D^{[n]}_{(n+1)b_n - 1} \right).$$

Now we note that for all $t \geq 0$, we have $\sigma_n^-(t) = X_{t-}^{(n)}$ and $\sigma_\infty^-(t) = X_{t^-}$. As a consequence, using (2.10), we get

$$\sigma_n^-(b_n) \xrightarrow[n\to\infty]{} X_{b^-} \qquad \text{a.s.}$$

The same relation for $\tilde{\sigma}_n^-$ and $\tilde{X}^{(n)}$, and the fact that $\tilde{b}_n$ converges a.s. to $\tilde{b} := 1 - b - l(b)$, show that

$$\sigma_n^+(b_n) = \tilde{\sigma}_n^-(\tilde{b}_n) \xrightarrow[n\to\infty]{} \tilde{X}_{\tilde{b}^-} \qquad \text{a.s.}$$

Thus $\sigma_n(b_n)$ converges a.s. to $\sigma_\infty^-(b) + \tilde{\sigma}_\infty^-(\tilde{b})$. To show that this quantity is equal to $\sigma_\infty(b)$, we introduce the "truncated" sums $\sigma_{n,\varepsilon}(t)$, $\sigma_{n,\varepsilon}^+(t)$, $\sigma_{n,\varepsilon}^-(t)$, obtained by taking into account only the $s \in (0, t)$ such that $X_{s-}^{(n)} < I_{st}^{(n)}$ and $\Delta X_s^{(n)} > \varepsilon$. For all $n \in \mathbb{N} \cup \{\infty\}$, these quantities are finite sums. Therefore, the a.s. convergence (2.10) implies that for all $\varepsilon > 0$,

$$\sigma_{\infty,\varepsilon}^+(b) = \lim_{n\to\infty} \sigma_{n,\varepsilon}^+(b_n) = \lim_{n\to\infty} \tilde{\sigma}_{n,\varepsilon}^-(\tilde{b}_n) = \tilde{\sigma}_{\infty,\varepsilon}^-(\tilde{b}).$$

Thus $\sigma_{\infty,\varepsilon}(b) = \sigma_{\infty,\varepsilon}^-(b) + \tilde{\sigma}_{\infty,\varepsilon}^-(\tilde{b})$. By letting $\varepsilon \to 0$, we get $\sigma_\infty(b) = \sigma_\infty^-(b) + \tilde{\sigma}_\infty^-(\tilde{b})$. □

We now come back to the proof of Proposition 2.5.

Proof of Proposition 2.5. For all $n \in \mathbb{N} \cup \{\infty\}$, we add edge-lengths to the discrete tree $\mathcal{R}_n(k)$ by letting

$$\ell_n(\{u, v\}) = d_n(u, v) \quad \text{if } n \in \mathbb{N},$$
$$\ell_\infty(\{u, v\}) = d(u, v),$$

for every edge $\{u, v\}$. Let $\mathcal{R}_n'(t)$ denote the resulting tree with edge-lengths. We now write $\mathcal{R}_n(k, t)$ for the tree $\mathcal{R}_n'(t)$ endowed with point processes of marks on its edges and vertices, defined as follows:

— The marks on the vertices of $\mathcal{R}_n(k)$ appear at the same time as the marks on the corresponding vertices of $\mathcal{T}_n$.
— Each edge receives a mark at its midpoint at the first time when a vertex v of $\mathcal{T}_n$ such that $v \in e$ is marked in $\mathcal{T}_n$.

For each n, these two point processes are independent, and their rates are the following:
— Each vertex $v \in V_n(k)$ is marked at rate $L_n(v)$, independently of the other vertices.
— For each edge e of $\mathcal{R}_n(k)$, letting b, b' denote the points of $B_n(k)$ corresponding to e^-, e^+ (as explained in the proof of Lemma 2.7), the edge e is marked at rate $\Sigma L_n(e)$, independently of the other edges, with

$$\begin{aligned}\Sigma L_n(e) &= \sum_{v \in V(\mathcal{T}_n)\cap e} L_n(v)\\ &= \sigma_n(b') - \sigma_n(b) + \frac{n}{a_n^2}\left(H^{(n)}_{(b')^-} - H^{(n)}_{b^-}\right) - L_n(e^-)\end{aligned}$$

if $n \in \mathbb{N}$, and

$$\Sigma L_\infty(e) = \Sigma L(e) = \sum_{v \in V(\mathcal{T})\cap e} L(v) = \sigma_\infty(b') - \sigma_\infty(b) - L(e^-).$$

Now Lemmas 2.7 and 2.8 show that $L_n(v)$ and $\Sigma L_n(e)$ converge to $L(v)$ and $\Sigma L(e)$ (respectively) as $n \to \infty$. Therefore we have the convergence

$$\left(\frac{a_n}{n}\mathcal{R}_n(k,t), t \geq 0\right) \xrightarrow[n\to\infty]{(d)} \left(\mathcal{R}_\infty(k,t), t \geq 0\right), \tag{2.11}$$

where $(a_n/n)\cdot\mathcal{R}_n(k,t)$ and $\mathcal{R}_\infty(k,t)$ can be seen as random variables in $\mathbb{T} \times (\mathbb{R}_+ \cup \{-1\})^{\mathbb{N}} \times \{-1,0,1\}^{\mathbb{N}^2}$: for example,

$$(a_n/n)\cdot\mathcal{R}_n(k,t) = \left(\mathcal{R}_n(k), (l_i)_{i\geq 1}, (\delta_V(i,t))_{i\geq 0}, (\delta_E(i,t))_{i \geq 1}\right),$$

where

$$l_i = \begin{cases} (a_n/n)\cdot \ell(e_i(\mathcal{R}_n(k))) & \text{if } i < N_n(k)\\ -1 & \text{if } i \geq N_n(k)\end{cases}$$

$$\delta_V(i,t) = \begin{cases} 1 & \text{if } i < N_n(k) \text{ and the vertex } v_i(\mathcal{R}_n(k))\\ & \qquad \text{has been marked before time } t\\ 0 & \text{if } i < N_n(k) \text{ and the vertex } v_i(\mathcal{R}_n(k))\\ & \qquad \text{has not been marked before time } t\\ -1 & \text{if } i \geq N_n(k)\end{cases}$$

$$\delta_E(i,t) = \begin{cases} 1 & \text{if } i < N_n(k) \text{ and the edge } e_i(\mathcal{R}_n(k))\\ & \qquad \text{has been marked before time } t\\ 0 & \text{if } i < N_n(k) \text{ and the edge } e_i(\mathcal{R}_n(k))\\ & \qquad \text{has not been marked before time } t\\ -1 & \text{if } i \geq N_n(k)\end{cases}$$

(recall that $N_n(k)$ is the number of vertices of $\mathcal{R}_n(k)$). Note that we could keep working under (2.10) to get an a.s. convergence, but this is no longer necessary.

The rest of the proof goes as in [19]. For every $i \in \mathbb{N}$, we let $\eta_n(k,i,t)$ denote the number of vertices among $\xi_n(1), \ldots, \xi_n(k)$ in the component of $\mathcal{R}_n(k)$ containing $\xi_n(i)$ at time t. Similarly, denote by $\eta_\infty(k,i,t)$ the number of vertices among $\xi(1),\ldots,\xi(k)$ in the component of $\mathcal{R}_\infty(k)$ containing $\xi(i)$ at time t. It follows from (2.11) that we have the joint convergences

$$\frac{a_n}{n}\mathcal{T}_n \xrightarrow{(d)} \mathcal{T}$$
$$(\eta_n(k,i,t))_{t\geq 0, i\in\mathbb{N}} \xrightarrow{(d)} (\eta_\infty(k,i,t))_{t\geq 0, i\in\mathbb{N}}$$
$$(\tau_n(i,j))_{i,j\in\mathbb{N}} \xrightarrow{(d)} (\tau(i,j))_{i,j\in\mathbb{N}}.$$

Besides, the law of large numbers gives that for each $i\in\mathbb{N}$ and $t\geq 0$,

$$\frac{1}{k}\eta_\infty(k,i,t) \xrightarrow[n\to\infty]{} \mu_{\xi(i)}(t) \qquad \text{a.s.}$$

Thus, for every fixed integer l and times $0\leq t_1\leq\ldots\leq t_l$, we can construct a sequence $k_n\to\infty$ sufficiently slowly, such that

$$\left(\frac{1}{k_n}\eta_n(k_n,i,t_j)\right)_{i,j\in\{1,\ldots,l\}} \xrightarrow{(d)} (\mu_{\xi(i)}(t_j))_{i,j\in\{1,\ldots,l\}},$$

or equivalently (see [9, Lemma 11])

$$\left(\mu_{n,\xi_n(i)}(t_j)\right)_{i,j\in\{1,\ldots,l\}} \xrightarrow{(d)} (\mu_{\xi(i)}(t_j))_{i,j\in\{1,\ldots,l\}},$$

both holding jointly with the preceding convergences. This entails the proposition. □

2.3 Upper bound for the expected component mass

To get the convergence of $(\mathcal{T}_n, \mathrm{Cut}_\mathrm{v}(\mathcal{T}_n))$, we will finally need to control the quantities

$$\mathbb{E}\left[\int_{2^l}^\infty \mu_{n,\xi_n}(t)dt\right],$$

where ξ_n is a uniform random integer in $\{1,\ldots,n\}$. Our main goal is to show that these quantities converge to 0 as l tends to ∞, uniformly in n, as stated in Corollary 2.15.

To this end, we will sometimes work under the size-biased measure GW^*, defined as follows. We recall that a pointed tree is a pair (T,v), where T is a rooted planar tree and v is a vertex of T. The measure GW^* is the sigma-finite measure such that, for every pointed tree (T,v),

$$GW^*(T,v) = \mathbb{P}\left(\mathbf{T}=T\right),$$

where $\mathbf{T}$ is a Galton–Watson tree with offspring distribution ν. We let $\mathbb{E}^*$ denote the expectation under this "law." In particular, the conditional law GW^* given $|V(T)|=n+1$ is well-defined, and corresponds to the distribution of a pair $(\mathcal{T}_n, v)$ where given $\mathcal{T}_n$, v is a uniform random vertex of $\mathcal{T}_n$. Hereafter, T will denote a ν-Galton–Watson tree, whose expectation will either be taken under the unbiased law or under a conditioned version of the law GW^*. Recall that we only consider values of n such that $P_n = \mathbb{P}\left(|V(T)|=n+1\right)\neq 0$.

For all $m,n\in\mathbb{N}$ such that $m\leq n$ and $P_m\neq 0$, for all $t\in\mathbb{R}_+$, we define

$$E_{m,n}(t) = \frac{1}{m}\mathbb{E}\left[\sum_{e\in E(\mathcal{T}_m)}\exp\left(-\sum_{u\in[\![\rho_m,e^-]\!]_V}\deg(u,\mathcal{T}_m)\frac{t}{a_n}\right)\right], \tag{2.12}$$

and $E_n(t)=E_{n,n}(t)$. Equivalently, we can write

$$E_{m,n}(t) = \frac{1}{m}\mathbb{E}^*\left[\sum_{e\in E(T)}\exp\left(-\sum_{u\in[\![\rho(T),e^-]\!]_V}\deg(u,T)\frac{t}{a_n}\right)\,\middle|\, |V(T)|=m+1\right].$$

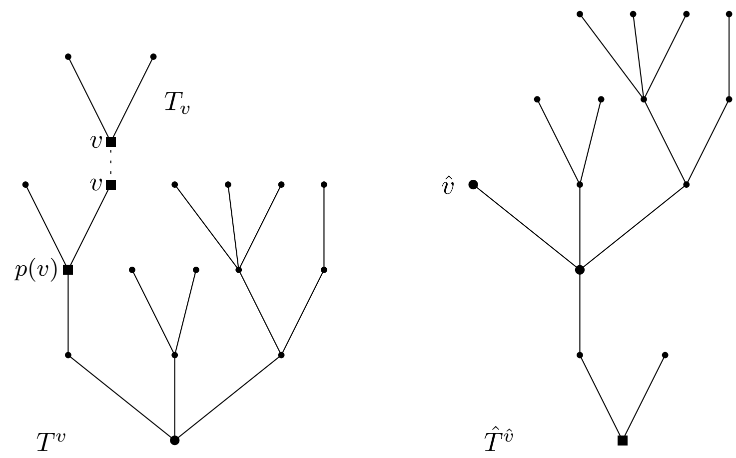

Figure 2.3 – The trees T_v, T^v and $\hat{T}^{\hat{v}}$ obtained from a pointed tree (T, v).

For all $m < n$, we also use the notation

$$P^*_{m,n} := \mathbb{P}^* \left(|V(T_v)| = m + 1 | \, |V(T)| = n + 1 \right),$$

where T_v denotes the tree formed by v and its descendants. Our first step is to show the following:

Lemma 2.9. *Let ξ_n be a uniform random edge of $\mathcal{T}_n$. Using the previous notation, we have*

$$\mathbb{E}\left[\mu_{n,\xi_n}(t)\right] \leq \frac{1}{n} e^{-t/a_n} + 2 \left(E_n(t) + \sum_{\substack{m=1 \\ P_m \neq 0}}^{n-1} P^*_{m,n} \frac{m}{n} E_{m,n}(t) \right). \tag{2.13}$$

The proof of this lemma will use Proposition 2.10 below. Let us first introduce some notation. For all $v \in V(T)$, we let T^v be the subtree obtained by deleting all the strict descendants of v in T, and as before, T_v be the tree formed by v and its descendants. We define a new tree $\hat{T}^{\hat{v}}$, constructed by taking T^v and modifying it as follows:

— we remove the edge $e(v)$ between v and $p(v)$;

— we add a new child $\hat{v}$ to the root, and let $\hat{e}_{\hat{v}}$ denote the edge between $\hat{v}$ and the root;

— we reroot the tree at $p(v)$.

An example of this construction is given in Figure 2.3. Note that we have natural bijective correspondences between $V(T)$, $(V(T^v) \setminus \{v\}) \sqcup V(T_v)$ and $(V(\hat{T}^{\hat{v}}) \setminus \{\hat{v}\}) \sqcup V(T_v)$, and between $E(T)$, $E(T^v) \sqcup E(T_v)$ and $E(\hat{T}^{\hat{v}}) \sqcup E(T_v)$. Furthermore, one can easily check that for all $u \in V(\hat{T}^{\hat{v}}) \setminus \{\hat{v}\}$, we have $\deg(u, \hat{T}^{\hat{v}}) = \deg(u, T)$, and for all $u \in V(T_v)$, $\deg(u, T_v) = \deg(u, T)$.

This transformation is the same as in [19, p.21], except that we work with rooted trees instead of planted trees. In our case, adding the edge $\hat{e}_{\hat{v}}$ and deleting $e(v)$ mimics the existence of a base edge. Thus, we can use Proposition 2 of [19]:

Proposition 2.10. *Under GW^*, $(\hat{T}^{\hat{v}}, T_v)$ and (T^v, T_v) have the same "law," and the trees T^v and T_v are independent, with T_v being a Galton–Watson tree.*

Proof of Lemma 2.9. In this proof, we identify ξ_n with the edge e_{ξ_n}, to make notation easier. We first note that for each edge $e \in E(\mathcal{T}_n)$, e belongs to the component $\mathcal{T}_{n,\xi_n}(t)$ if and only if no vertex on the path $[\![e^-, \xi_n^-]\!]_V$ has been removed at time t. Given $\mathcal{T}_n$ and ξ_n, this happens with probability

$$\exp\left(- \sum_{u \in [\![e^-, \xi_n^-]\!]_V} \deg u \cdot \frac{t}{a_n} \right)$$

(for any vertex u, at time t, u has been deleted from the initial tree with probability $1 - \exp(-\deg u \cdot t/a_n))$. Thus,

$$\mathbb{E}\left[n\mu_{n,\xi_n}\right] = \mathbb{E}\left[\sum_{e\in E(\mathcal{T}_n)} \mathbf{1}_{e\in\mathcal{T}_{n,\xi_n}(t)}\right] = \mathbb{E}\left[\sum_{e\in E(\mathcal{T}_n)} \exp\left(-\sum_{u\in[\![e^-,\xi_n^-]\!]_V} \deg u\cdot\frac{t}{a_n}\right)\right].$$

Since the edge ξ_n is chosen uniformly in $E(\mathcal{T}_n)$, this yields

$$\begin{aligned}\mathbb{E}\left[n\mu_{n,\xi_n}\right] &= \frac{1}{n}\mathbb{E}\left[\sum_{e,\xi\in E(\mathcal{T}_n)} \exp\left(-\sum_{u\in[\![e^-,\xi^-]\!]_V} \deg u\,\frac{t}{a_n}\right)\right]\\ &= \frac{1}{n}\mathbb{E}\left[\sum_{v\in V(\mathcal{T}_n)} \mathbf{1}_{v\neq\rho(\mathcal{T}_n)} \sum_{e\in E(\mathcal{T}_n)} \exp\left(-\sum_{u\in[\![e^-,p(v)]\!]_V} \deg u\,\frac{t}{a_n}\right)\right],\end{aligned}$$

where $p(v)$ denotes the parent of vertex v. Hence, calling $A_n(T)$ the event $\{|V(T)| = n+1\}$,

$$\mathbb{E}\left[n\mu_{n,\xi_n}\right] = \frac{n+1}{n}\mathbb{E}^*\left[\mathbf{1}_{v\neq\rho(T)}\sum_{e\in E(T)} \exp\left(-\sum_{u\in[\![e^-,p(v)]\!]_V} \deg u\,\frac{t}{a_n}\right)\middle| A_n(T)\right].$$

Distinguishing the cases for which $e \in E(T_v), e \in E(T^v)\setminus\{e(v)\}$ and $e = e(v)$, we split this quantity into three terms:

$$\mathbb{E}\left[n\mu_{n,\xi_n}\right] = \left(1+\frac{1}{n}\right)(\Sigma_v + \Sigma^v + \varepsilon_v), \tag{2.14}$$

where

$$\begin{aligned}\Sigma_v &= \mathbb{E}^*\left[\mathbf{1}_{v\neq\rho(T)}\sum_{e\in E(T_v)} \exp\left(-\sum_{u\in[\![e^-,v]\!]_V} (\deg(u,T_v)+\deg p(v))\,\frac{t}{a_n}\right)\middle| A_n(T)\right],\\ \Sigma^v &= \mathbb{E}^*\left[\mathbf{1}_{v\neq\rho(T)}\sum_{e\in E(T^v)\setminus\{e(v)\}} \exp\left(-\sum_{u\in[\![e^-,p(v)]\!]_V} \deg(u,T^v)\frac{t}{a_n}\right)\middle| A_n(T)\right],\end{aligned}$$

and

$$\varepsilon_v = \mathbb{E}^*\left[\mathbf{1}_{v\neq\rho(T)}\exp\left(-\deg p(v)\frac{t}{a_n}\right)\middle| A_n(T)\right].$$

For the first term, we have

$$\Sigma_v \leq \mathbb{E}^*\left[\mathbf{1}_{v\neq\rho(T)}\sum_{e\in E(T_v)} \exp\left(-\sum_{u\in[\![\rho(T_v),e^-]\!]_V} \deg(u,T_v)\frac{t}{a_n}\right)\middle| A_n(T)\right].$$

Since $|V(T)| = |V(T_v)| + |V(T^v)| - 1$, this gives

$$\Sigma_v \leq \sum_{\substack{m=1\\ P_m\neq 0}}^{n-1} P^*_{m,n}\mathbb{E}^*\left[\sum_{e\in E(T_v)} \exp\left(-\sum_{u\in[\![\rho(T_v),e^-]\!]_V} \deg(u,T_v)\frac{t}{a_n}\right)\middle| \begin{array}{l}|V(T_v)| = m+1,\\ |V(T^v)| = n-m+1\end{array}\right]$$

($m = n$ would correspond to the case where $v = \rho(T)$, and $m = 0$ to the case where $E(T_v) = \emptyset$). Proposition 2.10 gives that the trees T_v and T^v are independent, with T_v being a Galton–Watson tree. Hence

$$\begin{aligned}\Sigma_v &\leq \sum_{\substack{m=1\\P_m\neq 0}}^{n-1} P^*_{m,n}\mathbb{E}^*\left[\sum_{e\in E(T)} \exp\left(-\sum_{u\in[\![\rho(T),e^-]\!]_V} \deg(u,T)\frac{t}{a_n}\right)\middle| A_m(T)\right] \\ &\leq \sum_{\substack{m=1\\P_m\neq 0}}^{n-1} P^*_{m,n} m E_{m,n}(t). \end{aligned} \tag{2.15}$$

For the second term, we use the correspondence between $E(T^v) \setminus \{e(v)\}$ and $E(\hat{T}^{\hat{v}}) \setminus \{\hat{e}_{\hat{v}}\}$, and the fact that $\rho(\hat{T}^{\hat{v}}) = p(v)$:

$$\Sigma^v = \mathbb{E}^*\left[\mathbf{1}_{v\neq\rho(T)} \sum_{e\in E(\hat{T}^{\hat{v}})\setminus\{\hat{e}_{\hat{v}}\}} \exp\left(-\sum_{u\in[\![\rho(\hat{T}^{\hat{v}}),e^-]\!]_V} \deg(u,\hat{T}^{\hat{v}})\frac{t}{a_n}\right)\middle| A_n(T)\right].$$

This gives

$$\Sigma^v \leq \mathbb{E}^*\left[\sum_{e\in E(\hat{T}^{\hat{v}})} \exp\left(-\sum_{u\in[\![\rho(\hat{T}^{\hat{v}}),e^-]\!]_V} \deg(u,\hat{T}^{\hat{v}})\frac{t}{a_n}\right)\middle| A_n(T)\right].$$

Using the fact that T^v and $\hat{T}^{\hat{v}}$ have the same law under GW^*, we get

$$\Sigma^v \leq \mathbb{E}^*\left[\sum_{e\in E(T^v)} \exp\left(-\sum_{u\in[\![\rho(T^v),e^-]\!]_V} \deg(u,T^v)\frac{t}{a_n}\right)\middle| A_n(T)\right].$$

Seeing $E(T^v)$ as a subset of $E(T)$, we can write

$$\Sigma^v \leq \mathbb{E}^*\left[\sum_{e\in E(T)} \exp\left(-\sum_{u\in[\![\rho(T),e^-]\!]_V} \deg(u,T)\frac{t}{a_n}\right)\middle| A_n(T)\right] = nE_n(t). \tag{2.16}$$

For the third term, we simply notice that

$$\varepsilon_v \leq \frac{n}{n+1}e^{-t/a_n}. \tag{2.17}$$

Putting together (2.15), (2.16) and (2.17) into (2.14), we finally get

$$\mathbb{E}\left[n\mu_{n,\xi_n}(t)\right] \leq e^{-t/a_n} + \left(1+\frac{1}{n}\right)\left(nE_n(t) + \sum_{\substack{m=1\\P_m\neq 0}}^{n-1} P^*_{m,n} m E_{m,n}(t)\right).$$

Thus

$$\mathbb{E}\left[\mu_{n,\xi_n}(t)\right] \leq \frac{1}{n}e^{-t/a_n} + \left(1+\frac{1}{n}\right)\left(E_n(t) + \sum_{\substack{m=1\\P_m\neq 0}}^{n-1} P^*_{m,n} \frac{m}{n} E_{m,n}(t)\right).$$

□

Next, we compute $E_{m,n}(t)$. To this end, we introduce two new independent sequences of i.i.d. variables:

— $(\hat{Z}_i)_{i\geq 1}$ with law $\hat{\nu}$, where $\hat{\nu}$ is the size-biased version of ν;
— $(N_i)_{i\geq 1}$, with same law as the number of vertices of a Galton–Watson tree with offspring distribution ν.

For all $k, h \in \mathbb{N}$, we also write

$$\hat{S}_h = \sum_{i=1}^{h} \hat{Z}_i \quad \text{and} \quad Y_k = \sum_{i=1}^{k} N_i.$$

Lemma 2.11. *For every $m, n \in \mathbb{N}$ such that $m \leq n$ and $P_m \neq 0$, one has*

$$E_{m,n}(t) = \frac{1}{mP_m} \sum_{1\leq h\leq k\leq m} e^{-kt/a_n} \mathbb{P}(\hat{S}_h = k) \mathbb{P}\left(Y_{k-h+1} = m - h + 1\right). \tag{2.18}$$

Proof. We first note that relation (2.12) can be written otherwise, using the one-to-one correspondence $e \mapsto e^+$ between $E(T)$ and $V(T) \setminus \{\rho(T)\}$:

$$E_{m,n}(t) = \frac{1}{m}\mathbb{E}\left[\sum_{v\in V(T)\setminus\rho(T)} \exp\left(-\sum_{u\in[\![\rho(T),p(v)]\!]_V} \deg(u,T)\frac{t}{a_n}\right) \middle| |E(T)| = m\right].$$

We thus have

$$\begin{aligned} E_{m,n}(t) &= \frac{1}{mP_m}\mathbb{E}\left[\sum_{v\in V(T)\setminus\rho(T)} \exp\left(-\sum_{u\in[\![\rho(T),p(v)]\!]_V} \deg(u,T)\frac{t}{a_n}\right), |E(T)| = m\right] \\ &= \frac{1}{mP_m}\mathbb{E}^*\left[\mathbf{1}_{v\neq\rho(T)} \exp\left(-\sum_{u\in[\![\rho(T),p(v)]\!]_V} \deg(u,T)\frac{t}{a_n}\right), |E(T)| = m\right]. \end{aligned}$$

We now use the following description of a typical pointed tree (T, v) under GW^* (see the proof of Proposition 2 of [19] and [46]):
— The "law" under GW^* of the distance $h(v)$ of the pointed vertex v to the root is the counting measure on $\mathbb{N} \cup \{0\}$.
— Conditionally on $h(v) = h$, the subtrees T_v and T^v are independent, with T_v being a Galton–Watson tree with offspring distribution ν, and T^v having GW_h^* law, which can be described as follows. T^v has a distinguished branch $B = \{u_1 = \rho(T^v), u_2, \ldots, u_{h+1} = v\}$ of length h. Every vertex of T^v has an offspring that is distributed independently of the other vertices, with offspring distribution ν for the vertices in $V(T^v) \setminus B$, $\hat{\nu}$ for the vertices $u_1, \ldots, u_h$, and u_{h+1} having no descendants. The tree T^v can thus be constructed inductively from the root u_1, by choosing the ith vertex u_i of the distinguished branch uniformly at random from the children of u_{i-1}.

In this representation, conditionally on having $h(v) = h$, $[\![\rho(T), p(v)]\!]_V$ equals $\{u_1, \ldots, u_h\}$ and, for every $i \in \{1, \ldots, h\}$,

$$\deg(u_i, T) = \hat{Z}_i.$$

Besides, the total number of vertices of T is the sum of the number of vertices h of $B \setminus \{v\}$, of $|V(T_v)|$, and of the $|V(T_u)|$ for u such that $p(u) \in B \setminus \{v\}$ and $u \notin B$. There are $\sum_{i=1}^{h} (\hat{Z}_i - 1)$ such trees T_u. Hence under GW^*:

$$|E(T)| = |V(T)| - 1 \overset{(d)}{=} Y_{\sum_{i=1}^{h}(\hat{Z}_i - 1)+1} + h - 1.$$

Thus

$$E_{m,n}(t) = \frac{1}{mP_m} \sum_{1 \le h} \mathbb{E}\left[\exp\left(-\sum_{i=1}^{h} \hat{Z}_i \frac{t}{a_n}\right), Y_{\sum_{i=1}^{h} \hat{Z}_i - h+1} = m-h+1\right]$$
$$= \frac{1}{mP_m} \sum_{1 \le h \le k \le m} e^{-kt/a_n} \mathbb{P}(\hat{S}_h = k) \mathbb{P}\left(Y_{k-h+1} = m-h+1\right).$$

□

We now compute upper bounds for the terms $\mathbb{P}\left(Y_{k-h+1} = m-h+1\right)$, $\mathbb{P}(\hat{S}_h = k)$ and $(mP_m)^{-1}$.

Upper bound for $\mathbb{P}\left(Y_{k-h+1} = m-h+1\right)$

Recalling the notation of Section 2.2.2, we have

$$\mathbb{P}\left(Y_k = n\right) = \mathbb{P}\left(W_n = -k \text{ and}, \forall p < n,\ W_p > -k\right)$$
$$= \frac{k}{n} \mathbb{P}\left(W_n = -k\right).$$

The second equality is given by the cyclic lemma (see [55, Lemma 6.1]). We will now use the fact, given by Theorem 2.2, that

$$\lim_{n\to\infty} \sup_{k\in\mathbb{N}} \left| a_n \mathbb{P}\left(W_n = -k\right) - p_1^{(\alpha)}\left(-\frac{k}{a_n}\right)\right| = 0. \tag{2.19}$$

For all $s, x \in (0, \infty)$, we have

$$x p_s^{(\alpha)}(-x) = s q_x^{(1/\alpha)}(s).$$

(see, e.g., [13, Corollary VII.1.3]). Taking $s = 1$ and $x = k/a_n$, this gives

$$\frac{k}{a_n} p_1^{(\alpha)}\left(-\frac{k}{a_n}\right) = q_{k/a_n}^{(1/\alpha)}(1).$$

Thus

$$n\mathbb{P}\left(Y_n = k\right) - q_{k/a_n}^{(1/\alpha)}(1) = \frac{k}{a_n}\left(a_n \mathbb{P}\left(W_n = -k\right) - p_1^{(\alpha)}\left(-\frac{k}{a_n}\right)\right),$$

and we get

$$\mathbb{P}\left(Y_k = n\right) \le \frac{1}{n}\left(\left|n\mathbb{P}\left(Y_n = k\right) - q_{k/a_n}^{(1/\alpha)}(1)\right| + q_{k/a_n}^{(1/\alpha)}(1)\right)$$
$$\le \frac{k}{na_n}\left(\left|a_n \mathbb{P}\left(W_n = -k\right) - p_1^{(\alpha)}\left(-\frac{k}{a_n}\right)\right| + p_1^{(\alpha)}\left(-\frac{k}{a_n}\right)\right).$$

Since $p_1^{(\alpha)}$ is bounded and (2.19) holds, there exists a constant $M \in (0, \infty)$ such that, for all $k, n \in \mathbb{N}$,

$$\mathbb{P}\left(Y_k = n\right) \le \frac{k}{na_n} M.$$

Thus, we have the following upper bound:

$$\mathbb{P}\left(Y_{k-h+1} = m-h+1\right) \le \frac{k-h+1}{(m-h+1)a_{m-h+1}} M. \tag{2.20}$$

Upper bound for $\mathbb{P}(\hat{S}_h = k)$

We use Theorem 2.2 for the i.i.d. variables $(\hat{Z}_i)_{i\in\mathbb{N}}$. Let $\hat{A} \in R_{\alpha-1}$ be an increasing function given by (i), such that

$$\mathbb{P}\left(\hat{Z}_1 > r\right) \sim \frac{1}{\hat{A}(r)},$$

and $\hat{a}$ be the inverse function of $\hat{A}$. Then

$$\lim_{h\to\infty} \sup_{k\in\mathbb{N}} \left| \hat{a}_h \mathbb{P}\left(\hat{S}_h = k\right) - q_1^{(\alpha-1)}\left(\frac{k}{\hat{a}_h}\right)\right| = 0.$$

Using the fact that $q_1^{(\alpha-1)}$ is bounded, and writing

$$\mathbb{P}\left(\hat{S}_h = k\right) \leq \frac{1}{\hat{a}_h}\left(\left|\hat{a}_h \mathbb{P}\left(\hat{S}_h = k\right) - q_1^{(\alpha-1)}\left(\frac{k}{\hat{a}_h}\right)\right| + q_1^{(\alpha-1)}\left(\frac{k}{\hat{a}_h}\right)\right),$$

we get the existence of a constant $M' \in (0,\infty)$ such that, for all $h, k \in \mathbb{N}$,

$$\mathbb{P}\left(\hat{S}_h = k\right) \leq \frac{M'}{\hat{a}_h}. \tag{2.21}$$

Furthermore, when h is small enough, we have a better bound for $\mathbb{P}(\hat{S}_h = k)$:

Lemma 2.12. *Using the previous notation, if hypothesis* (2.1) *holds, then there exist constants* B, C *such that for all* $k \in \mathbb{N}$*, for all* h *such that* $k/\hat{a}_h \geq B$*,*

$$\mathbb{P}\left(\hat{S}_h = k\right) \leq C\frac{h}{k\hat{A}(k)}.$$

This result is an adaptation of a theorem by Doney [29]. The main ideas of the proof, which is rather technical, will be given in the Appendix.

Besides, using the fact that A is regularly varying and an Abel transformation of $\mathbb{P}(\hat{Z} > r)$, we get that

$$\frac{1}{\hat{A}(r)} \sim \frac{\alpha r}{A(r)} \quad \text{as } r \to \infty. \tag{2.22}$$

Upper bound for $(mP_m)^{-1}$

We have

$$P_m = \mathbb{P}\left(|E(\mathcal{T})| = m\right) \sim \frac{p_1^{(\alpha)}(0)}{ma_m}$$

(this is a straightforward consequence of the cyclic lemma and the local limit theorem). This gives the existence of a constant $K \in (0,\infty)$ which verifies, for all m such that $P_m \neq 0$,

$$\frac{1}{mP_m} \leq Ka_m. \tag{2.23}$$

Before coming back to the proof of Corollary 2.15, we give another useful result on regularly varying functions.

Lemma 2.13. *Fix* $\beta \in (0,\infty)$*. Let* f *be a positive increasing function in* R_β *on* $\mathbb{R}_+$*, and* x_0 *a positive constant. For every* $\delta \in (0,\beta)$*, there exists a constant* $C_\delta \in (0,\infty)$ *such that, for all* $x' \geq x \geq x_0$*,*

$$C_\delta^{-1}\left(\frac{x'}{x}\right)^{\beta-\delta} \leq \frac{f(x')}{f(x)} \leq C_\delta\left(\frac{x'}{x}\right)^{\beta+\delta}.$$

This result is a consequence of the Potter bounds (see, e.g., Theorem 1.5.6. of Bingham *et al.* [20]). In particular, it implies that for all x bounded away from 0, for all $z \geq 1$,

$$C_\delta^{-1} z^{\beta-\delta} \leq \frac{f(xz)}{f(x)} \leq C_\delta z^{\beta+\delta}, \tag{2.24}$$

and likewise, for all $x \in (0, \infty)$, $z \leq 1$ such that xz is bounded away from 0,

$$C_\delta^{-1} z^{\beta+\delta} \leq \frac{f(xz)}{f(x)} = \frac{f(xz)}{f(xzz^{-1})} \leq C_\delta z^{\beta-\delta}. \tag{2.25}$$

We can finally state the following.

Lemma 2.14. *We have*

$$\lim_{l\to\infty} \sup_{n\in\mathbb{N}} \int_{2^l}^{\infty} E_n(t)dt = 0, \tag{2.26}$$

and

$$\lim_{l\to\infty} \sup_{n\in\mathbb{N}} \sup_{\substack{1\leq m\leq n \\ P_m\neq 0}} \int_{2^l}^{\infty} \frac{m}{n} E_{m,n}(t)dt = 0.$$

Proof. For every $n, l \in \mathbb{N}$, we let

$$I_{n,l} = \int_{2^l}^{\infty} E_n(t)dt.$$

Putting together (2.18) and (2.23), we have

$$E_n(t) \leq K a_n \sum_{k=1}^{n} \sum_{h=1}^{k} e^{-kt/a_n} \mathbb{P}\left(\hat{S}_h = k\right) \mathbb{P}\left(Y_{k-h+1} = n-h+1\right).$$

This yields

$$I_{n,l} \leq K a_n^2 \sum_{k=1}^{n} \sum_{h=1}^{k} \frac{1}{k} e^{-2^l k/a_n} \mathbb{P}\left(\hat{S}_h = k\right) \mathbb{P}\left(Y_{k-h+1} = n-h+1\right).$$

Writing $h(n,k) = \hat{A}(k/B) \wedge \lfloor n/2 \rfloor$ and $h'(n,k) = k \wedge \lfloor n/2 \rfloor$, we split this sum into three parts:

$$I_{n,l}^1 = a_n^2 \sum_{k=1}^{n} \sum_{h=1}^{h(n,k)} \frac{1}{k} e^{-2^l k/a_n} \mathbb{P}\left(\hat{S}_h = k\right) \mathbb{P}\left(Y_{k-h+1} = n-h+1\right)$$

$$I_{n,l}^2 = a_n^2 \sum_{k=1}^{n} \sum_{h=h(n,k)+1}^{h'(n,k)} \frac{1}{k} e^{-2^l k/a_n} \mathbb{P}\left(\hat{S}_h = k\right) \mathbb{P}\left(Y_{k-h+1} = n-h+1\right)$$

$$I_{n,l}^3 = a_n^2 \sum_{k=1}^{n} \sum_{h=h'(n,k)+1}^{k} \frac{1}{k} e^{-2^l k/a_n} \mathbb{P}\left(\hat{S}_h = k\right) \mathbb{P}\left(Y_{k-h+1} = n-h+1\right).$$

Our first goal is to show that, for $i = 1, 2, 3$,

$$\lim_{l\to\infty} \sup_{n\in\mathbb{N}} I_{n,l}^i = 0.$$

Let us first examine $I^1_{n,l}$. Since a is increasing, the upper bound (2.20) gives, for $n-h+1 \geq n/2$,

$$\mathbb{P}\left(Y_{k-h+1} = n-h+1\right) \leq M\frac{k-h+1}{(n-k+1)a_{n-k+1}} \\ \leq 2M\frac{k}{na_{n/2}}. \tag{2.27}$$

Thus, we have

$$I^1_{n,l} \leq 2M\frac{a_n^2}{na_{n/2}}\sum_{k=1}^{n} e^{-2^l k/a_n}\sum_{h=1}^{h(n,k)} \mathbb{P}\left(\hat{S}_h = k\right).$$

Turning the first sum into an integral, and using the substitution $y' = y/a_n$, we get

$$\begin{aligned} I^1_{n,l} &\leq 2M\frac{a_n^2}{na_{n/2}}\int_1^\infty dy e^{-2^l\lfloor y\rfloor/a_n}\left(\sum_{h=1}^{h(n,\lfloor y\rfloor)} \mathbb{P}\left(\hat{S}_h = \lfloor y\rfloor\right)\right) \\ &= 2M\frac{a_n^3}{na_{n/2}}\int_{1/a_n}^\infty dy e^{-2^l\lfloor a_n y\rfloor/a_n}\left(\sum_{h=1}^{h(n,\lfloor a_n y\rfloor)} \mathbb{P}\left(\hat{S}_h = \lfloor a_n y\rfloor\right)\right). \end{aligned}$$

Since $\hat{a}$ is increasing, for all $h \leq h(n,k)$, we have $\hat{a}_h \leq k/B$. Therefore, Lemma 2.12 gives

$$\mathbb{P}\left(\hat{S}_h = k\right) \leq C\frac{h}{k\hat{A}(k)}.$$

This yields

$$\begin{aligned} I^1_{n,l} &\leq 2CM\frac{a_n^3}{na_{n/2}}\int_{1/a_n}^\infty dy e^{-2^l\lfloor a_n y\rfloor/a_n}\left(\sum_{h=1}^{h(n,\lfloor a_n y\rfloor)} \frac{h}{a_n y\hat{A}(a_n y)}\right) \\ &\leq 2CM\frac{a_n^3}{na_{n/2}}\int_{1/a_n}^\infty dy e^{-2^l\lfloor a_n y\rfloor/a_n}\left(\frac{\hat{A}(\lfloor a_n y\rfloor/B)^2}{\lfloor a_n y\rfloor\hat{A}(\lfloor a_n y\rfloor)}\right). \end{aligned}$$

We fix $\delta \in (0,(\alpha-1)\wedge(2-\alpha))$. Since $\hat{A}$ is regularly varying with index $\alpha-1$, for all $y \geq 1/a_n$, we have

$$\frac{\hat{A}(\lfloor a_n y\rfloor/B)}{\hat{A}(\lfloor a_n y\rfloor)} \leq \frac{C_\delta^{-1}}{B^{\alpha-1-\delta}}$$

(we can use (2.24) because $\lfloor a_n y\rfloor/B \geq 1/B$ for all $y \in (1/a_n,\infty), n \in \mathbb{N}$). As a consequence, there exists a positive constant K_1 such that

$$I^1_{n,l} \leq K_1\frac{a_n^3}{na_{n/2}}\int_{1/a_n}^\infty dy e^{-2^l\lfloor a_n y\rfloor/a_n}\left(\frac{\hat{A}(\lfloor a_n y\rfloor)}{\lfloor a_n y\rfloor}\right) = K_1 J_{n,l}.$$

Therefore, it suffices to show that

$$\lim_{l\to\infty}\sup_{n\in\mathbb{N}} J_{n,l} = 0. \tag{2.28}$$

To this end, we use the upper bounds 2.24 and 2.25, with $x = a_n$ and $y = \lfloor a_n y\rfloor/a_n$ (x and xy being respectively greater than a_0 and 1):

$$\frac{\hat{A}(\lfloor a_n y\rfloor)}{\hat{A}(a_n)} \leq C_\delta\left(\left(\frac{\lfloor a_n y\rfloor}{a_n}\right)^{\alpha-1+\delta} \vee \left(\frac{\lfloor a_n y\rfloor}{a_n}\right)^{\alpha-1-\delta}\right).$$

Thus

$$J_{n,l} \leq \frac{a_n^2 \hat{A}(a_n)}{n a_{n/2}} \int_{1/a_n}^{\infty} dy e^{-2^l \lfloor a_n y \rfloor / a_n} \left(\left(\frac{a_n}{\lfloor a_n y \rfloor} \right)^{2-\alpha-\delta} \vee \left(\frac{a_n}{\lfloor a_n y \rfloor} \right)^{2-\alpha+\delta} \right).$$

Using the fact that $\lfloor a_n y \rfloor \geq a_n y - 1$, and the change of variable $y' = y - 1/a_n$, we get

$$J_{n,l} \leq \frac{a_n^2 \hat{A}(a_n)}{n a_{n/2}} \int_{0}^{\infty} dy e^{-2^l y} \left(\frac{1}{y^{2-\alpha-\delta}} \vee \frac{1}{y^{2-\alpha+\delta}} \right).$$

Now (2.22) gives that $\hat{A}(a_n)/n = \hat{A}(a_n)/A(a_n) \sim 1/\alpha a_n$, so we have

$$\frac{a_n^2 \hat{A}(a_n)}{n a_{n/2}} \sim \frac{a_n}{\alpha a_{n/2}}.$$

Since a is regularly varying with index $1/\alpha$, the right-hand term has a finite limit as n goes to infinity. Therefore $a_n^2 \hat{A}(a_n)/n a_{n/2}$ is bounded uniformly in n. Hence there exists a constant $K \in (0, \infty)$ such that

$$\sup_{n \in \mathbb{N}} J_{n,l} \leq K \int_{0}^{\infty} dy e^{-2^l y} \left(\frac{1}{y^{2-\alpha-\delta}} \vee \frac{1}{y^{2-\alpha+\delta}} \right).$$

This yields (2.28) by taking the limit as l goes to infinity.

For the second part, we can still use (2.27). As in the first step, we get

$$I_{n,l}^2 \leq 2M \frac{a_n^3}{n a_{n/2}} \int_{1/a_n}^{\infty} dy e^{-2^l \lfloor a_n y \rfloor / a_n} \left(\sum_{h=h(n, \lfloor a_n y \rfloor)+1}^{h'(n, \lfloor a_n y \rfloor)} \mathbb{P}\left(\hat{S}_h = \lfloor a_n y \rfloor \right) \right).$$

Since the sum is null if $\hat{A}(\lfloor a_n y \rfloor / B) > \lfloor n/2 \rfloor$, we have

$$I_{n,l}^2 \leq 2M \frac{a_n^3}{n a_{n/2}} \int_{1/a_n}^{\infty} dy e^{-2^l \lfloor a_n y \rfloor / a_n} \left(\sum_{h=\hat{A}(\lfloor a_n y \rfloor / B)+1}^{\infty} \mathbb{P}\left(\hat{S}_h = \lfloor a_n y \rfloor \right) \right).$$

We now turn the remaining sum into an integral:

$$I_{n,l}^2 \leq 2M \frac{a_n^3}{n a_{n/2}} \int_{1/a_n}^{\infty} dy e^{-2^l \lfloor a_n y \rfloor / a_n} \int_{\hat{A}(\lfloor a_n y \rfloor / B)}^{\infty} dx \, \mathbb{P}\left(\hat{S}_{\lfloor x+1 \rfloor} = \lfloor a_n y \rfloor \right).$$

Using the change of variable $x' = \hat{A}(\lfloor a_n y \rfloor / B) x$ and the upper bound (2.21), this gives

$$I_{n,l}^2 \leq 2MM' \frac{a_n^3}{n a_{n/2}} \int_{1/a_n}^{\infty} dy e^{-2^l \lfloor a_n y \rfloor / a_n} \int_{1}^{\infty} dx \frac{\hat{A}(\lfloor a_n y \rfloor / B)}{\hat{a}\left(\lfloor \hat{A}(\lfloor a_n y \rfloor / B) x + 1 \rfloor \right)}.$$

Since $\hat{a}$ is increasing, for all x, y, we have

$$\hat{a}\left(\lfloor \hat{A}(\lfloor a_n y \rfloor / B) x + 1 \rfloor \right) \geq \hat{a}\left(\hat{A}(\lfloor a_n y \rfloor / B) x \right).$$

Fix $\delta \in (0, 1/(\alpha-1) - 1)$. Inequality (2.24) then gives, for all $x \geq 1$, $y \geq 1/a_n$,

$$\begin{aligned} \hat{a}\left(\lfloor \hat{A}(\lfloor a_n y \rfloor / B) x + 1 \rfloor \right) &\geq c_\delta^{-1} \hat{a}\left(\hat{A}(\lfloor a_n y \rfloor / B) \right) x^{1/(\alpha-1)-\delta} \\ &= c_\delta^{-1} \frac{\lfloor a_n y \rfloor}{B} x^{1/(\alpha-1)-\delta}. \end{aligned}$$

Thus there exist constants $K_2, K_2' \in (0,\infty)$ such that

$$I_{n,l}^2 \leq K_2 \frac{a_n^3}{n a_{n/2}} \int_{1/a_n}^{\infty} dy e^{-2^l \lfloor a_n y \rfloor / a_n} \frac{\hat{A}(\lfloor a_n y \rfloor / B)}{\lfloor a_n y \rfloor} \int_1^{\infty} \frac{dx}{x^{1/(\alpha-1)-\delta}} = K_2' J_{n,l},$$

and (2.28) also gives the conclusion.

For the third part, since the terms with indices $k \leq \lfloor n/2 \rfloor$ are null, we simply use the bounds $\mathbb{P}\left(Y_{k-h+1} = n-h+1\right) \leq 1$ and $\mathbb{P}(\hat{S}_h = k) \leq 1$:

$$\begin{aligned} I_{n,l}^3 &\leq a_n^2 \sum_{k=\lfloor n/2 \rfloor+1}^{n} \sum_{h=1}^{k} \frac{1}{k} e^{-2^l k / a_n} \\ &\leq a_n^2 e^{-n 2^l / 2 a_n} \sum_{k=\lfloor n/2 \rfloor +1}^{n} 1 \\ &\leq n a_n^2 e^{-n 2^l / 2 a_n}. \end{aligned}$$

This quantity tends to 0 as l goes to infinity, uniformly in n. Indeed, for any $\kappa > 0$, the function $g_\kappa : x \mapsto x^\kappa e^{-x}$ is bounded by a constant G_κ, hence

$$I_{n,l}^3 \leq G_\kappa \frac{2^\kappa a_n^{2+\kappa}}{n^{\kappa-1}} \cdot 2^{-l\kappa}.$$

For any $\varepsilon > 0$, there exists a constant C_ε such that $a_n \leq C_\varepsilon n^{1/\alpha+\varepsilon}$ for all $n \in \mathbb{N}$. Therefore, the quantity $a_n^{2+\kappa}/n^{\kappa-1}$ is bounded as soon as $\kappa > (2+\alpha)/(\alpha-1)$. This completes the proof of (2.26).

For the second limit, we note that (2.18) yields

$$\int_{2^l}^{\infty} E_{m,n}(t) dt = \frac{a_n}{a_m} \int_{2^l}^{\infty} E_m(t) dt,$$

for all $m \leq n$ such that $P_m \neq 0$. Thus

$$\sup_{n \in \mathbb{N}} \sup_{\substack{1 \leq m \leq n \\ P_m \neq 0}} \int_{2^l}^{\infty} \frac{m}{n} E_{m,n}(t) dt = \sup_{n \in \mathbb{N}} \sup_{\substack{1 \leq m \leq n \\ P_m \neq 0}} \frac{m a_n}{n a_m} I_{m,l}.$$

As a consequence, it is enough to show that $m a_n / n a_m$ is bounded over $\{(m,n) \in \mathbb{N}^2 : m \leq n\}$. Now,

$$\begin{aligned} \sup\left\{\frac{m a_n}{n a_m} : m,n \in \mathbb{N},\ m \leq n\right\} &\leq \sup\left\{\frac{m a_{\lambda m}}{\lambda m a_m} : m \in \mathbb{N},\ \lambda \in (1,\infty)\right\} \\ &\leq \sup\left\{\frac{a_{\lambda m}}{\lambda a_m} : m \in \mathbb{N},\ \lambda \in (1,\infty)\right\}. \end{aligned}$$

Fix $\delta \in (0, 1-1/\alpha)$. Since a is a positive increasing function in $R_{1/\alpha}$, Lemma 2.13 shows the existence of a constant such that, for all $m \in \mathbb{N}$, $\lambda \in (1,\infty)$,

$$\frac{a_{\lambda m}}{a_m} \leq C_\delta \lambda^{1/\alpha+\delta}.$$

Hence, for all $\lambda \in (1,\infty)$,

$$\sup_{m \in \mathbb{N}} \frac{a_{\lambda m}}{\lambda a_m} \leq C_\delta \lambda^{1/\alpha+\delta-1} \leq C_\delta.$$

□

Key estimates for the proof of Theorem 1.3

We conclude this section by giving two consequences of Lemma 2.14 which will be used in the proof of Theorem 1.3.

Corollary 2.15. *It holds that*

$$\lim_{l\to\infty}\sup_{n\in\mathbb{N}}\mathbb{E}\left[\int_{2^l}^{\infty}\mu_{n,\xi_n}(t)dt\right]=0.$$

Proof. Using (2.13), we get

$$\begin{aligned}\sup_{n\in\mathbb{N}}\mathbb{E}\left[\int_{2^l}^{\infty}\mu_{n,\xi_n}(t)dt\right]\le \sup_{n\in\mathbb{N}}\frac{a_n}{n}e^{-2^l/a_n}+2\sup_{n\in\mathbb{N}}\int_{2^l}^{\infty}E_n(t)dt\\ +2\sup_{n\in\mathbb{N}}\sup_{1\le m\le n}\int_{2^l}^{\infty}\frac{m}{n}E_{m,n}(t)dt.\end{aligned}$$

Lemma 2.14 shows that the last two terms tend to 0 as l goes to infinity. For the first term, we use again the fact that for any $\kappa>0$, the function $g_\kappa : x\mapsto x^\kappa e^{-x}$ is bounded by a constant G_κ. Hence, for all $n\in\mathbb{N}$,

$$\frac{a_n}{n}e^{-2^l/a_n}\le G_\kappa\frac{a_n^{\kappa+1}}{n}\cdot 2^{-\kappa l}.$$

Taking $\kappa<\alpha-1$, we get that $a_n^{\kappa+1}/n$ is bounded, which completes the proof. □

Corollary 2.16. *There exists a constant C such that, for all $n\in\mathbb{N}$,*

$$\mathbb{E}\left[\delta_n'(0,\xi_n)\right]\le C.$$

Proof. Recalling the definition of δ_n', we get

$$\mathbb{E}\left[\delta_n'(0,\xi_n)\right]=\mathbb{E}\left[\int_0^{\infty}\mu_{n,\xi_n}(t)dt\right].$$

Now the upper bound (2.13) gives

$$\begin{aligned}\mathbb{E}\left[\delta_n'(0,\xi_n)\right]&\le 1+\mathbb{E}\left[\int_1^{\infty}\mu_{n,\xi_n}(t)dt\right]\\ &\le 1+\frac{a_n}{n}e^{-1/a_n}+2\int_1^{\infty}E_n(t)dt+2\sup_{1\le m\le n}\int_1^{\infty}\frac{m}{n}E_{m,n}(t)dt.\end{aligned}$$

The second term is bounded as $n\to\infty$. Recall from the proof of Lemma 2.14 that

$$\begin{aligned}\int_1^{\infty}E_n(t)dt=I_{n,0}&\le I_{n,0}^1+I_{n,0}^2+I_{n,0}^3\\ &\le (K_1+K_2')J_{n,0}+I_{n,0}^3.\end{aligned}$$

Moreover, we have seen that for any $\delta>0$, there exists a constant K such that

$$\sup_{n\in\mathbb{N}}J_{n,0}\le K\int_0^{\infty}dye^{-y}\left(\frac{1}{y^{2-\alpha-\delta}}\wedge\frac{1}{y^{2-\alpha+\delta}}\right)<\infty,$$

and

$$I_{n,0}^3\le 2na_n^2e^{-n/a_n}$$

is bounded as $n\to\infty$. Since we have seen at the end of the proof of Lemma 2.9 that there exists a constant K' such that for all $n\in\mathbb{N}$, $m\le n$ such that $P_m\neq 0$,

$$\int_1^{\infty}\frac{m}{n}E_{m,n}(t)dt\le K'\int_1^{\infty}E_m(t)dt,$$

this implies the corollary. □

3 Proof of Theorem 1.3

3.1 Identity in law between $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T})$ and $\mathcal{T}$

In this section, we show that the semi-infinite matrices of the mutual distance of uniformly sampled points in $\mathcal{T}$ and $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T})$ have the same law. This justifies the existence of $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T})$, as explained in Section 1.2, and shows the identity in law between $\mathcal{T}$ and $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T})$. The structure of the proof will be similar to that of Lemma 4 in [19]. Precise descriptions of the fragmentation processes we consider can be found in [49] and [50].

Recall that $(\xi(i))_{i\in\mathbb{N}}$ is a sequence of i.i.d. random variables in $\mathcal{T}$, with law μ, and $\xi(0)=0$. Since the law of $\mathcal{T}$ is invariant under uniform rerooting (see, e.g., [33, Proposition 4.8]), and the definition of δ does not depend on the choice of the root of $\mathcal{T}$, we may assume that $\xi(1)=\rho$.

Proposition 3.1. *It holds that*

$$(\delta(\xi(i),\xi(j)))_{i,j\geq 0} \overset{(d)}{=} (d(\xi(i+1),\xi(j+1)))_{i,j\geq 0}\,.$$

Proof. Here, it is convenient to work on fragmentation processes taking values in the set of the partitions of $\mathbb{N}$.

First, we introduce a process Π which corresponds to our fragmentation of $\mathcal{T}$ by saying that $i,j\in\mathbb{N}$ belong to the same block of $\Pi(t)$ if and only if the path $[\![\xi(i),\xi(j)]\!]_V$ does not intersect the set $\{b_k : k\in I, t_k\leq t\}$ of the points marked before time t. For every $i\in\mathbb{N}$, we let $B_i(t)$ be the block of the partition $\Pi(t)$ containing i. Note that the partitions $\Pi(t)$ are exchangeable, which justifies the existence of the asymptotic frequencies $\lambda(B_i(t))$ of the blocks $B_i(t)$, where

$$\lambda(B)=\lim_{n\to\infty}\frac{1}{n}\left|B\cap\{1,\ldots,n\}\right|.$$

Then we define

$$\sigma_i(t)=\inf\left\{u\geq 0 : \int_0^u \lambda(B_i(s))ds > t\right\}.$$

We use σ_i as a time-change, letting $\Pi'(t)$ be the partition whose blocks are the sets $B_i(\sigma_i(t))$ for $i\in\mathbb{N}$. Note that this is possible because $B_i(\sigma_i(t))$ and $B_j(\sigma_j(t))$ are either equal or disjoint.

We define a second fragmentation Γ, which results from cutting the stable tree $\mathcal{T}$ at its heights. For every $x,y\in\mathcal{T}$, we let $x\wedge y$ denote the branch-point between x and y, i.e. the unique point such that $[\![\rho,x\wedge y]\!]_V=[\![\rho,x]\!]_V\cap[\![\rho,y]\!]_V$. With this notation, we say that $i,j\in\mathbb{N}$ belong to the same block of $\Gamma(t)$ if and only if $d(\rho,\xi(i+1)\wedge\xi(j+1))>t$.

Then we have the following link between the two fragmentations:

Lemma 3.2. *The fragmentation processes Π' and Γ have the same law.*

Proof. Miermont has shown in [50, Theorem 1] that the process Π is a self-similar fragmentation with index $1/\alpha$, erosion coefficient 0 and dislocation measure Δ_α known explicitly. Applying Theorem 3.3 in [16], we get that the time-changed fragmentation Π' is still self-similar, with index $1/\alpha-1$, erosion coefficient 0 and the same dislocation measure Δ_α. Now the process Γ is also self-similar, with the same characteristics as Π' (see [49, Proposition 1, Theorem 1]). Thus Γ and Π' have the same law. $\square$

Using the law of large numbers, we note that $\lambda(B_i(s))=\mu_{\xi(i)}(s)$ almost surely. As a consequence, $\sigma_i(t)=\infty$ for $t=\int_0^\infty\lambda(B_i(s))ds=\delta(0,\xi(i))$, which means that $\delta(0,\xi(i))$ can be seen as the first time when the singleton $\{i\}$ is a block of Π'. Recalling that $d(\rho,\xi(i+1))=d(\xi(1),\xi(i+1))$ is the first time when $\{i\}$ is a block of Γ, we get

$$(\delta(0,\xi(i)))_{i\geq 1}\overset{(d)}{=}(d(\xi(1),\xi(i+1)))_{i\geq 1}\,. \tag{2.29}$$

Similarly, for any $i \neq j \in \mathbb{N}$,

$$\delta(0, \xi(i) \wedge \xi(j)) = \frac{1}{2}(\delta(0, \xi(i)) + \delta(0, \xi(j)) - \delta(\xi(i), \xi(j)))$$
$$= \int_0^{\tau(i,j)} \lambda(B_i(s))ds,$$

where $\tau(i,j)$ denotes the first time when a mark appears on the segment $[\![\xi(i), \xi(j)]\!]_V$. Thus $\delta(0, \xi(i) \wedge \xi(j))$ is the first time when the blocks containing i and j are separated in Π'. In terms of the fragmentation Γ, this corresponds to $d(\rho, \xi(i+1) \wedge \xi(j+1))$. Hence

$$(\delta(0, \xi(i) \wedge \xi(j)))_{i,j\geq 1} \stackrel{(d)}{=} (d(\xi(1), \xi(i+1) \wedge \xi(j+1)))_{i,j\geq 1},$$

and this holds jointly with (2.29). This entails the proposition. □

3.2 Weak convergence

We first establish the convergence for the cut-tree $\mathrm{Cut}'_{\mathrm{v}}(\mathcal{T}_n)$ endowed with the modified distance δ'_n, as defined in Section 2.1.

Proposition 3.3. *There is the joint convergence*

$$\left(\frac{a_n}{n}\mathcal{T}_n, \mathrm{Cut}'_{\mathrm{v}}(\mathcal{T}_n)\right) \xrightarrow[n\to\infty]{(d)} (\mathcal{T}, \mathrm{Cut}_{\mathrm{v}}(\mathcal{T}))$$

in $\mathbb{M} \times \mathbb{M}$.

Proof. Proposition 2.5 shows that for every fixed integer l, there is the joint convergence

$$\frac{a_n}{n}\mathcal{T}_n \xrightarrow[n\to\infty]{(d)} \mathcal{T}$$
$$\left(2^{-l}\sum_{j=1}^{4^l} \mu_{n,\xi_n(i)}\left(j2^{-l}\right)\right)_{i\in\mathbb{N}} \xrightarrow[n\to\infty]{(d)} \left(2^{-l}\sum_{j=1}^{4^l} \mu_{\xi(i)}\left(j2^{-l}\right)\right)_{i\in\mathbb{N}}.$$

Let

$$\Delta_{n,l}(i) = \mathbb{E}\left[\left|\int_0^\infty \mu_{n,\xi_n(i)}(t)dt - 2^{-l}\sum_{j=1}^{4^l} \mu_{n,\xi_n(i)}\left(j2^{-l}\right)\right|\right].$$

For any nonincreasing function $f : \mathbb{R}_+ \to [0,1]$, we have the upper bound:

$$\left|\int_0^\infty f(t)dt - 2^{-l}\sum_{j=1}^{4^l} f\left(j2^{-l}\right)\right| \leq 2^{-l} + \int_{2^l}^\infty f(t)dt. \tag{2.30}$$

Applying this inequality to $\mu_{n,\xi_n(i)}$ yields

$$\Delta_{n,l}(i) \leq 2^{-l} + \mathbb{E}\left[\int_{2^l}^\infty \mu_{n,\xi_n}(t)dt\right].$$

Corollary 2.15 now shows that

$$\lim_{l\to\infty} \sup_{n\in\mathbb{N}} \Delta_{n,l}(i) = 0,$$

and $\Delta_{n,l}(i)$ does not depend on i. Besides, Proposition 3.1 shows that

$$\delta(0, \xi(i)) = \int_0^\infty \mu_{\xi(i)}(t)dt - --$$

has the same law as $d(0,\xi(i))$ and, therefore, has finite mean. As a consequence

$$\mathbb{E}\left[\left|\int_0^\infty \mu_{\xi(i)}(t)dt - 2^{-l}\sum_{j=1}^{4^l}\mu_{\xi(i)}\left(j2^{-l}\right)\right|\right] \le 2^{-l} + \mathbb{E}\left[\int_{2^l}^\infty \mu_{\xi(i)}(t)dt\right] \xrightarrow[l\to\infty]{} 0,$$

and the left-hand side does not depend on i. We conclude that

$$\left(\delta'_n(0,\xi_n(i))\right)_{i\in\mathbb{N}} \xrightarrow[n\to\infty]{(d)} \left(\delta(0,\xi(i))\right)_{i\in\mathbb{N}},$$

jointly with $(a_n/n)\cdot\mathcal{T}_n \xrightarrow{(d)} \mathcal{T}$.

Using in addition the convergence of the $\tau_n(i,j)$ shown in Proposition 2.5, a similar argument shows that the preceding convergences also hold jointly with

$$\left(\delta'_n(\xi_n(i),\xi_n(j))\right)_{i,j\in\mathbb{N}} \xrightarrow[n\to\infty]{(d)} \left(\delta(\xi(i),\xi(j))\right)_{i,j\in\mathbb{N}}.$$

This entails the proposition. □

The convergence stated in Theorem 1.3 now follows immediately. Indeed, Lemma 2.1 and Corollary 2.16 show that

$$\mathbb{E}\left[\left|\frac{a_n}{n}\delta_n(i,j) - \delta'_n(i,j)\right|^2\right] \le \frac{2Ca_n}{n}$$

for all $i,j\ge 0$ (recalling that $\xi_n(0)=0$). Thus the preceding proposition gives the joint convergence

$$\left(\frac{a_n}{n}\mathcal{T}_n, \frac{a_n}{n}\operatorname{Cut}_{\mathrm{v}}(\mathcal{T}_n)\right) \xrightarrow[n\to\infty]{(d)} \left(\mathcal{T}, \operatorname{Cut}_{\mathrm{v}}(\mathcal{T})\right).$$

4 The finite variance case

In this section, we assume that the offspring distribution ν of the Galton–Watson trees $\mathcal{T}_n$ has finite variance σ^2. Theorem 23 of [7] shows that $(\sigma/\sqrt{n})\cdot\mathcal{T}_n$ converges to the Brownian tree $\mathcal{T}^{br}$. More precisely, still using the three processes described in Section 2.2.2 to encode the trees $\mathcal{T}_n$, the joint convergence stated in Theorem 2.3 holds with $a_n=\sigma\sqrt{n}$, and limit processes defined by $X_t=B_t$ and $H_t=2B_t$ for all $t\in[0,1]$. (Recall that B denotes the excursion of length 1 of the standard Brownian motion.) Note that the normalization of X is not exactly the same as the one we used for the stable tree, since the Laplace transform of a standard Brownian motion B' is $\mathbb{E}[e^{-\lambda B'_t}]=e^{\lambda^2 t/2}$. The fact that the height process H is equal to $2X$ can be seen from the definition of H as a local time, as explained in [32, Section 1.2].

Given these results, the proof of Theorem 1.4 follows the same structure as that of the main theorem. We first note that the results on the modified distance, introduced in Section 2.1, still hold. In the next two sections, we will see that we also have analogues for Proposition 2.5, and Corollaries 2.15 and 2.16.

4.1 Convergence of the component masses

We use the same notation as in Section 2.2. Recall in particular that $\mu_{n,\xi_n(i)}$ denotes the mass of the component $\mathcal{T}_{n,\xi_n(i)}(t)$, and that $\tau_n(i,j)$ denotes the first time when the components $\mathcal{T}_{n,\xi_n(i)}(t)$ and $\mathcal{T}_{n,\xi_n(j)}(t)$ become disjoint. To simplify, we drop the superscript *br* for the quantities associated to the Brownian tree (e.g., the mass-measure, the mass of a component, etc.), keeping the notation we used in the case of the stable tree. Our first step is to prove the following result.

Proposition 4.1. *As $n \to \infty$, we have the following weak convergences*

$$\frac{\sigma}{\sqrt{n}} \mathcal{T}_n \xrightarrow{(d)} \mathcal{T}^{br}$$

$$(\tau_n(i,j))_{i,j\geq 0} \xrightarrow{(d)} \left(\left(1 + \frac{1}{\sigma^2}\right)^{-1} \tau(i,j) \right)_{i,j\geq 0}$$

$$\left(\mu_{n,\xi_n(i)}(t)\right)_{i\geq 0, t\geq 0} \xrightarrow{(d)} \left(\mu_{\xi(i)} \left(\left(1 + \frac{1}{\sigma^2}\right) t \right) \right)_{i\geq 0, t\geq 0},$$

where the three hold jointly.

We begin by showing the same kind of property as in Lemma 2.4. For all $n \in \mathbb{N}$, we let $\tilde{X}^{(n)}$ and $\tilde{C}^{(n)}$ denote the rescaled Lukasiewicz path and contour function of the symmetrized tree $\tilde{\mathcal{T}}_n$.

Lemma 4.2. *We have the joint convergence*

$$(X^{(n)}, C^{(n)}, \tilde{X}^{(n)}, \tilde{C}^{(n)}) \xrightarrow[n\to\infty]{(d)} (X, H, \tilde{X}, \tilde{H}),$$

where $\tilde{H}_t = H_{1-t}$ and $\tilde{X}_t = \tilde{H}_t/2$ for all $t \in [0,1]$.

Proof. Since $\mathcal{T}_n$ and $\tilde{\mathcal{T}}_n$ have the same law, $(\tilde{X}^{(n)}, \tilde{C}^{(n)})$ converges in distribution to a couple of processes having the same law as (X, H) in $\mathbb{D} \times \mathbb{D}$. Thus the sequence of the laws of the processes $(X^{(n)}, C^{(n)}, \tilde{X}^{(n)}, \tilde{C}^{(n)})$ is tight in $\mathbb{D}^4$. Up to extraction, we can assume that $(X^{(n)}, C^{(n)}, \tilde{X}^{(n)}, \tilde{C}^{(n)})$ converges in distribution to $(X, H, \tilde{X}, \tilde{H})$.

Fix $t \in [0,1]$. The definition of the contour function shows that for all $n \in \mathbb{N}$, we have $\tilde{C}_t^{(n)} = C_{1-t}^{(n)}$. Since H and $\tilde{H}$ are a.s. continuous, taking the limit yields $\tilde{H}_t = H_{1-t}$ almost surely. Besides, since (X, H) and $(\tilde{X}, \tilde{H})$ have the same law, we have $\tilde{X}_t = \tilde{H}_t/2$ a.s. for all $t \in [0,1]$.

These equalities also hold a.s., simultaneously for a countable number of times t, and the continuity of H, X, $\tilde{H}$ and $\tilde{X}$ give that a.s., they hold for all $t \in [0,1]$. This identifies uniquely the law of $(X, H, \tilde{X}, \tilde{H})$, hence the lemma. □

This lemma shows that we can still work in the setting of

$$\begin{cases} \left(X^{(n)}, \tilde{X}^{(n)}\right) \xrightarrow[n\to\infty]{} \left(X, \tilde{X}\right) & \text{a.s.} \\ \left(t_i^{(n)}, i \in \mathbb{N}\right) \xrightarrow[n\to\infty]{} (t_i, i \in \mathbb{N}) & \text{a.s.} \end{cases} \tag{2.31}$$

where $t_i^{(n)} = (\xi_n(i)+1)/(n+1)$ for all $n \in \mathbb{N}$, $i \geq 0$, and $(t_i, i \in \mathbb{N})$ is a sequence of independent uniform variables in $[0,1]$ such that $\xi(i) = p(t_i)$.

Recall the notation $\mathcal{R}_n(k)$ for the shape of the subtree of $\mathcal{T}_n$ (or $\mathcal{T}^{br}$ if $n = \infty$) spanned by the root and the vertices $\xi_n(1), \ldots, \xi_n(k)$ (or $\xi(1), \ldots, \xi(k)$ if $n = \infty$). We also keep the notation $L_n(v) = \deg(v, \mathcal{T}_n)/a_n$ for the rate at which a vertex v is deleted in $\mathcal{T}_n$ (if $n \in \mathbb{N}$), and

$$\sigma_n(t) = \sum_{\substack{0<s<t \\ X_{s-}^{(n)} < I_{st}^{(n)}}} \Delta X_s^{(n)} \qquad \forall t \in [0,1],$$

where $I_{s,t}^{(n)} = \inf_{s<u<t} X_u^{(n)}$, and $X^{(\infty)} = X$.

As in Section 2.2, we state two lemmas which allow us two control the rates at which the fragmentations happen on the vertices and the edges of $\mathcal{R}_n(k)$.

Lemma 4.3. *Fix $k \in \mathbb{N}$. Under (2.31), $\mathcal{R}_n(k)$ is a.s. constant for all n large enough (say $n \geq N$). Identifying the vertices of $\mathcal{R}_n(k)$ with $\mathcal{R}_\infty(k)$ for all $n \geq N$, we have the a.s. convergence*

$$L_n(v) \xrightarrow[n\to\infty]{} 0 \qquad \forall v \in V(\mathcal{R}_\infty(k)).$$

Proof. The proof is the same as that of Lemma 2.7. In particular, we get that if the $b^{(n,k)}$ are the times encoding the "same" vertex v of $R_n(k)$, for $n \geq N$, then we have the a.s. convergences

$$\begin{aligned} b^{(n,k)} &\xrightarrow[n\to\infty]{} b^{(\infty,k)} \\ X^{(n)}_{b^{(n,k)}} &\xrightarrow[n\to\infty]{} X_{b^{(\infty,k)}} \\ X^{(n)}_{(b^{(n,k)})^-} &\xrightarrow[n\to\infty]{} X_{(b^{(\infty,k)})^-}. \end{aligned}$$

Since X is now continuous, this yields

$$L_n(v) = \Delta X^{(n)}_{b^{(n,k)}} + \frac{1}{a_n} \xrightarrow[n\to\infty]{} \Delta X_{b^{(\infty,k)}} = 0.$$

□

Lemma 4.4. *Let $(b_n)_{n\geq 1} \in [0,1]^{\mathbb{N}}$ be a converging sequence in $[0,1]$, and let b denote its limit. Then*

$$\sigma_n(b_n) \xrightarrow[n\to\infty]{} H_b \quad a.s.$$

Proof. As in the proof of Lemma 2.8, for all $n \in \mathbb{N} \cup \{\infty\}$, we write $\sigma_n(t) = \sigma_n^-(t) + \sigma_n^+(t)$, where

$$\sigma_n^+(t) = \sum_{\substack{0<s<t \\ X^{(n)}_{s-}<I^{(n)}_{st}}} \left(X^{(n)}_s - I^{(n)}_{s,t} \right) \quad \text{and} \quad \sigma_n^-(t) = \sum_{\substack{0<s<t \\ X^{(n)}_{s-}<I^{(n)}_{st}}} \left(I^{(n)}_{s,t} - X^{(n)}_{s^-} \right).$$

For all $t \geq 0$, $n \in \mathbb{N}$, we have $\sigma_n^-(t) = X^{(n)}_{t^-}$. As a consequence, (2.31) gives

$$\sigma_n^-(b_n) \xrightarrow[n\to\infty]{} X_b \qquad \text{a.s.}$$

Besides, we still have $\sigma_n^+(b_n) = \tilde{\sigma}_n^-(\tilde{b}_n)$, with

$$\tilde{b}_n = 1 - b_n + \frac{1}{n+1}\left(1 + H^{[n]}_{(n+1)b_n-1} - D^{[n]}_{(n+1)b_n-1}\right).$$

Now

$$\tilde{b}_n \xrightarrow[n\to\infty]{} 1 - b - l(b),$$

where $l(b) = \inf\{s > b : X_s = X_b\} - b$. Using (2.31) again, we get

$$\sigma_n^+(b_n) \xrightarrow[n\to\infty]{} \tilde{X}_{1-b-l(b)} = X_{b+l(b)} = X_b \qquad \text{a.s.}$$

Thus we have the a.s. convergence

$$\sigma_n(b_n) \xrightarrow[n\to\infty]{} 2X_b = H_b.$$

□

We can now give the proof of Proposition 4.1.

Proof of Proposition 4.1. Fix $n \in \mathbb{N} \cup \{\infty\}$. As in the proof of Proposition 2.5, we write $\mathcal{R}_n(k,t)$ for the reduced tree with edge-lengths, endowed with point processes of marks on its edges and vertices such that:

— The marks on the vertices of $\mathcal{R}_n(k)$ appear at the same time as the marks on the corresponding vertices of $\mathcal{T}_n$.

— Each edge receives a mark at its midpoint at the first time when a vertex v of $\mathcal{T}_n$ such that $v \in e$ is marked in $\mathcal{T}_n$.

These two point processes are independent, and their rates are the following:

— If $n \in \mathbb{N}$, each vertex v of $\mathcal{R}_n(k)$ is marked at rate $L_n(v)$, independently of the other vertices. If $n = \infty$, there are no marks on the vertices.

— For each edge e of $\mathcal{R}_n(k)$, letting b, b' denote the points of $B_n(k)$ corresponding to e^-, e^+, the edge e is marked at rate $\Sigma L_n(e)$, independently of the other edges, with

$$\begin{aligned}\Sigma L_n(e) &= \sum_{v \in V(\mathcal{T}_n) \cap e} L_n(v) \\ &= \sigma_n(b') - \sigma_n(b) + \frac{n}{a_n^2}\left(H^{(n)}_{(b')^-} - H^{(n)}_{b^-}\right) - L_n(e^-)\end{aligned}$$

if $n \in \mathbb{N}$, and

$$\Sigma L_\infty(e) = H_{b'} - H_b.$$

We see from Lemmas 4.3 and 4.4 that $L_n(v)$ converges to 0 as $n \to \infty$, and that

$$\Sigma L_n(e) \xrightarrow[n\to\infty]{} \left(1 + \frac{1}{\sigma^2}\right) \Sigma L_\infty(e).$$

As a consequence, we have the convergence

$$\left(\frac{a_n}{n}\mathcal{R}_n(k,t), t \geq 0\right) \xrightarrow[n\to\infty]{(d)} \left(\mathcal{R}_\infty\left(k, \left(1 + \frac{1}{\sigma^2}\right) t\right), t \geq 0\right). \tag{2.32}$$

(As in the case $\alpha \in (1,2)$, $(a_n/n) \cdot \mathcal{R}_n(k,t)$ and $\mathcal{R}_\infty(k,t)$ can be seen as random variables in $\mathbb{T} \times (\mathbb{R}_+ \cup \{-1\})^{\mathbb{N}} \times \{-1,0,1\}^{\mathbb{N}^2}$.)

For all $i \in \mathbb{N}$, we let $\eta_n(k,i,t)$ denote the number of vertices among $\xi_n(1), \ldots, \xi_n(k)$ in the component of $\mathcal{R}_n(k)$ containing $\xi_n(i)$ at time t, and similarly $\eta_\infty(k,i,t)$ the number of vertices among $\xi(1), \ldots, \xi(k)$ in the component of $\mathcal{R}_\infty(k)$ containing $\xi(i)$ at time t. It follows from (2.32) that we have the joint convergences

$$\begin{aligned}\frac{a_n}{n}\mathcal{T}_n &\xrightarrow{(d)} \mathcal{T}^{br} \\ (\eta_n(k,i,t))_{t\geq 0, i \in \mathbb{N}} &\xrightarrow{(d)} \left(\eta_\infty\left(k, i, \left(1 + \frac{1}{\sigma^2}\right) t\right)\right)_{t \geq 0, i \in \mathbb{N}} \\ (\tau_n(i,j))_{i,j \in \mathbb{N}} &\xrightarrow{(d)} \left(\left(1 + \frac{1}{\sigma^2}\right)^{-1} \tau(i,j)\right)_{i,j \in \mathbb{N}}.\end{aligned}$$

The end of the proof is the same as for Proposition 2.5. $\square$

4.2 Upper bound for the expected component mass

The second step is to show that, as in Section 2.3, the following properties hold:

Lemma 4.5. *It holds that*

$$\lim_{l \to \infty} \sup_{n \in \mathbb{N}} \mathbb{E}\left[\int_{2^l}^{\infty} \mu_{n,\xi_n}(t) dt\right] = 0.$$

Besides, there exists a constant C such that, for all $n \in \mathbb{N}$,

$$\mathbb{E}\left[\delta'_n(0, \xi_n)\right] \leq C.$$

Proof. We use the fact that there exists a natural coupling between the edge-fragmentation and the vertex-fragmentation of $\mathcal{T}_n$. Indeed, both can be obtained by a deterministic procedure, given $\mathcal{T}_n$ and a uniform permutation $(i_1,\dots,i_n)$ of $\{1,\dots,n\}$. More precisely, in the edge-fragmentation, we delete the edge e_{i_k} at each step k, thus splitting $\mathcal{T}_n$ into at most two connected components, whereas in the vertex fragmentation, we delete all the edges such that $e^- = e^-_{i_k}$. Thus, at each step, the connected component containing a given edge e for the vertex-fragmentation is included in the component containing e for the edge-fragmentation.

Now consider the continuous-time versions of these fragmentations: each edge is marked independently with rate $a_n/n = \sigma/\sqrt{n}$ in our case, and $1/\sqrt{n}$ in [19]. We let $\mathcal{T}^E_{n,i}(t)$ and $\mathcal{T}^V_{n,i}(t)$ denote the connected components containing the edge e_i at time t, respectively for the edge-fragmentation and the vertex-fragmentation. Then the preceding remark shows that there exists a coupling such that $\mathcal{T}^V_{n,i}(t) \subset \mathcal{T}^E_{n,i}(\sigma t)$ a.s., and thus $\mu_n(\mathcal{T}^V_{n,i}(t)) \le \mu_n(\mathcal{T}^E_{n,i}(\sigma t))$ almost surely.

Lemma 3 and Corollary 1 of [19] show that the two announced properties hold for the case of the edge-fragmentation. Therefore, they also hold for the vertex-fragmentation. □

4.3 Proof of Theorem 1.4

As before, the proof of Theorem 1.4 now relies on showing a joint convergence for the rescaled versions of $\mathcal{T}_n$ and the modified cut-tree $\mathrm{Cut}_{\mathrm{v}}(\mathcal{T}_n)$:

$$\left(\frac{a_n}{n}\mathcal{T}_n, \left(1+\frac{1}{\sigma^2}\right)\mathrm{Cut}'_{\mathrm{v}}(\mathcal{T}_n)\right) \xrightarrow[n\to\infty]{(d)} \left(\mathcal{T}^{br}, \mathrm{Cut}(\mathcal{T}^{br})\right) \tag{2.33}$$

in $\mathbb{M}\times\mathbb{M}$. Indeed, Lemma 2.1 and the second part of Lemma 4.5 show that

$$\mathbb{E}\left[\left|\frac{a_n}{n}\delta_n(i,j) - \delta'_n(i,j)\right|^2\right] \le \frac{2Ca_n}{n}$$

for all $i,j\ge 0$. Thus (2.33) entails the joint convergence

$$\left(\frac{a_n}{n}\mathcal{T}_n, \frac{a_n}{n}\left(1+\frac{1}{\sigma^2}\right)\mathrm{Cut}_{\mathrm{v}}(\mathcal{T}_n)\right) \xrightarrow[n\to\infty]{(d)} \left(\mathcal{T}, \mathrm{Cut}_{\mathrm{v}}(\mathcal{T})\right).$$

Since $a_n = \sigma\sqrt{n}$, this gives Theorem 1.4.

Let us finally justify why (2.33) holds. Proposition 4.1 shows that for every fixed integer l, there is the joint convergence

$$\frac{a_n}{n}\mathcal{T}_n \xrightarrow[n\to\infty]{(d)} \mathcal{T}^{br}$$
$$\left(2^{-l}\sum_{j=1}^{4^l}\mu_{n,\xi_n(i)}\left(j2^{-l}\right)\right)_{i\in\mathbb{N}} \xrightarrow[n\to\infty]{(d)} \left(2^{-l}\sum_{j=1}^{4^l}\mu_{\xi(i)}\left(C_\sigma j2^{-l}\right)\right)_{i\in\mathbb{N}},$$

where $C_\sigma = 1+1/\sigma^2$. Using the upper bound (2.30) and the first part of Lemma 4.5, we get that

$$\lim_{l\to\infty}\sup_{n\in\mathbb{N}}\mathbb{E}\left[\left|\int_0^\infty \mu_{n,\xi_n(i)}(t)dt - 2^{-l}\sum_{j=1}^{4^l}\mu_{n,\xi_n(i)}\left(j2^{-l}\right)\right|\right] = 0,$$

and these expectations do not depend on i. Proposition 3.1 of [19] shows that $\delta(0,\xi(i))$ has the same law as $d(0,\xi(i))$ and, therefore, has finite mean. Thus,

$$\left|\int_0^\infty \mu_{\xi(i)}\left(C_\sigma t\right)dt - 2^{-l}\sum_{j=1}^{4^l}\mu_{\xi(i)}\left(C_\sigma j2^{-l}\right)\right| \le \underbrace{2^{-l} + \mathbb{E}\left[\int_{2^l}^\infty \mu_{\xi(i)}\left(C_\sigma t\right)dt\right]}_{\xrightarrow[l\to\infty]{}0},$$

and the left-hand side does not depend on i. Since

$$\int_0^\infty \mu_{\xi(i)}\left(C_\sigma t\right) dt = C_\sigma^{-1} \int_0^\infty \mu_{\xi(i)}(t) dt = C_\sigma^{-1}\delta(0,\xi(i)),$$

we conclude that

$$\left(C_\sigma \delta'_n(0,\xi_n(i))\right)_{i\in\mathbb{N}} \xrightarrow[n\to\infty]{(d)} \left(\delta(0,\xi(i))\right)_{i\in\mathbb{N}},$$

jointly with $(a_n/n)\cdot \mathcal{T}_n \xrightarrow{(d)} \mathcal{T}$. Using in addition the convergence of the $\tau_n(i,j)$ shown in Proposition 2.5, we see that the preceding convergences also hold jointly with

$$\left(C_\sigma \delta'_n(\xi_n(i),\xi_n(j))\right)_{i,j\in\mathbb{N}} \xrightarrow[n\to\infty]{(d)} \left(\delta(\xi(i),\xi(j))\right)_{i,j\in\mathbb{N}},$$

and this gives the convergence (2.33).

Appendix: Adaptation of Doney's result

We rephrase Lemma 2.12 using the notation of [29].

Lemma 4.6. *Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of i.i.d. variables in $\mathbb{N}\cup\{0\}$, whose law belongs to the domain of attraction of a stable law of index $\hat{\alpha}\in(0,1)$, and $S_n = X_1+\ldots+X_n$. We also let $A\in R_{\hat{\alpha}}$ be a positive increasing function such that*

$$\mathbb{P}\left(X>r\right)\sim\frac{1}{A(r)}, \tag{2.34}$$

and a the inverse function of A. Besides, we suppose that the additional hypothesis

$$\sup_{r\geq 1}\left(\frac{r\mathbb{P}(X=r)}{\mathbb{P}(X>r)}\right)<\infty \tag{2.35}$$

holds. Then there exists constants B, C such that for all $r\in\mathbb{N}$, for all n such that $r/a_n\geq B$,

$$\mathbb{P}\left(S_n=r\right)\leq C\frac{n}{rA(r)}.$$

This result is an adaptation of a theorem shown by Doney in [29], which gives an equivalent for $\mathbb{P}\left(S_n=r\right)$ as $n\to\infty$, uniformly in n such that $r/a_n\to\infty$, using the slightly stronger hypothesis

$$\mathbb{P}\left(X=r\right)\sim\frac{1}{rA(r)} \qquad \text{as } r\to\infty$$

instead of (2.35).

Sketch of the proof. The main idea is to split up $\mathbb{P}\left(S_n=r\right)$ into four terms, depending upon the values taken by $M_n=\max\left\{X_i : i=1,\ldots,n\right\}$ and $N_n=\left|\{m\leq n : X_m>z\}\right|$. More precisely, letting η and γ be constants in $(0,1)$, $w=r/a_n$ and $z=a_n w^\gamma$, we have

$$\mathbb{P}\left(S_n=r\right)=\sum_{i=0}^{3}\mathbb{P}\left(\{S_n=r\}\cap A_i\right),$$

where $A_i=\{M_n\leq\eta r, N_n=i\}$ for $i=0,1$, $A_2=\{M_n\leq\eta r, N_n\geq 2\}$ and $A_3=\{M_n>\eta r\}$. For our purposes, it is enough to show that there exists constants c_i such that

$$q_i:=\mathbb{P}\left(\{S_n=r\}\cap A_i\right)\leq c_i\frac{n}{rA(r)} \qquad \forall i\in\{0,1,2,3\}.$$

The constants γ and η are fixed, with conditions that will be given later (see the detailed version of the proof for explicit conditions). In the whole proof, we suppose that $w \geq B$, for B large enough (possibly depending on the values of η and γ). Note that hypotheses (2.34) and (2.35) imply the existence of a constant c such that

$$p_r = \mathbb{P}(X = r) \leq \frac{c}{rA(r)} \qquad \text{and} \qquad \overline{F}(r) = \mathbb{P}(X > r) \leq \frac{c}{A(r)}. \tag{2.36}$$

The first calculations of [29] show that we have the following inequalities:

$$\begin{aligned} q_3 &\leq n \sup_{l > \eta r} p_l \\ q_2 &\leq \frac{1}{2} n^2 \overline{F}(z) \sup_{l > z} p_l \\ q_1 &\leq n \mathbb{P}(M_{n-1} \leq z, S_{n-1} > (1-\eta) r) \sup_{l > z} p_l. \end{aligned}$$

We now use (2.36), and apply Lemma 2.13 for the regularly varying function A. The first inequality thus yields the existence of a constant c_3 which only depends on the value of η. Similarly, the second inequality gives the existence of c_2, provided γ is large enough (independently of B) and $B \geq 1$.

To get the existence of c_1, we first apply Lemma 2 of [29], which gives an upper bound for the quantity $\mathbb{P}(M_{n-1} \leq z, S_{n-1} > (1-\eta)r)$ provided z is large enough and $(1-\eta)r \geq z$. Since $a_1 w^\gamma \leq z \leq r/w^{1-\gamma}$, these conditions can be achieved by taking B large enough. The lemma gives

$$q_1 \leq c \frac{n}{zA(z)} \cdot \left(\frac{c' z}{(1-\eta) r} \right)^{(1-\eta) r / z},$$

where c' is a constant. Now, applying Lemma 2.13, we get the existence of a constant c'_1 such that

$$q_1 \leq c'_1 \frac{n}{rA(r)} \cdot w^\kappa,$$

where κ depends on the values of η, γ and B. For a given choice of η and γ, and for B large enough, κ is negative, hence the existence of c_1.

For q_0, getting the upper bound goes by first showing that we can work under the hypotheses $r \leq nz$ and $r \leq n a_n/2$ (instead of the hypotheses $n \to \infty$ and $r/n a_n \to 0$ of [29]). Indeed, if $r > nz$, then $q_0 = 0$, and if $r > n a_n/2$, another application of Lemma 2 of [29] and of Lemma 2.13 yields the result. The rest of the proof relies on replacing the X_i by truncated variables $\hat{X}_i$, and using an exponentially biased probability law. This last part is long and technical, but it is rather easy to check that each step still holds with our hypotheses, for B large enough and with an appropriate choice of η (independently of B). $\square$

Chapitre 3

Inverting the cut-tree transform

Les résultats de ce chapitre sont le fruit d'une collaboration avec Lougi Addario-Berry et Christina Goldschmidt. La version présentée ici sert de base à un article en cours d'écriture.

1 Introduction

The articles [19, 28] introduced a transformation of the Brownian and stable trees which, given a fragmentation of the initial tree $\mathcal{T}$, yields a new random continuous tree called the cut-tree of $\mathcal{T}$. This cut-tree $\mathrm{Cut}(\mathcal{T})$ describes the genealogy of the connected components created by the fragmentation, and is shown to have the same distribution as the initial tree.

In the present work, we give a way to rebuild the initial tree from its cut-tree. In the same way that the cut-tree $\mathrm{Cut}(\mathcal{T})$ depends on $\mathcal{T}$ and on the extra randomness of the fragmentation, our construction associates to a stable tree $\mathcal{C}$ a reconstructed tree $\mathrm{Rec}(\mathcal{C})$ which depends on $\mathcal{C}$ and on an extra randomness given by a *decoration* on $\mathcal{C}$. We show that for a well-chosen decoration, this transformation has the following properties:

- if $\mathcal{C}$ is a stable tree of index $\alpha \in (1, 2)$ or a Brownian tree, then $\mathrm{Rec}(\mathcal{C})$ has the same distribution as $\mathcal{C}$.
- if $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$ almost surely, then we have $\mathrm{Rec}(\mathcal{C}) = \mathcal{T}$ almost surely,

Note that the question of the reconstruction has been studied by Broutin and Wang [21] in the Brownian case. Our decoration is close in spirit to the extra randomness they use. However, we use a different approximation in our reconstruction; also note that our results are slightly stronger, in the sense that we identify the decoration for which we have the almost sure equality $\mathrm{Rec}(\mathrm{Cut}(\mathcal{T})) = \mathcal{T}$.

We first explain the underlying idea of the reconstruction, and in particular the role of the decoration, in the discrete case.

1.1 The discrete case

Let T be a finite tree with n vertices labelled $v_1, \ldots, v_n$. To simplify our notation, we suppose that T is rooted; however, the choice of the root will have no influence on our construction. We consider a fragmentation obtained by deleting a uniform random edge of T at each step, until all the edges are removed. Let us first explain how the cut-tree $\mathrm{Cut}(T)$ is built. This construction is similar to that of [19], but is not exactly the same, the main difference being that the leaves of $\mathrm{Cut}(T)$ correspond to the vertices of T instead of its edges.

Fix $r \in \{1, \ldots, n-1\}$. We let $e(r)$ be the edge that is deleted at step r, $\overline{T}(r)$ the forest obtained from T after the deletion of the r first edges, and for all $i \in \{1, \ldots, n\}$, we let $T(i, r)$ denote the connected component of $\overline{T}(r)$ containing the vertex v_i. We define an equivalence relation $\sim_r$ on $\{1, \ldots, n\}$ by saying that $i \sim_r i'$ if and only if $T(i, r) = T(i', r)$. The family of the equivalence classes (without repetition) of the relations $\sim_r$ for $r = 1, \ldots, N$ forms the set of the vertices of $\mathrm{Cut}(T)$. The initial block $\{1,\dots,n\}$ is seen as the root, and the leaves of $\mathrm{Cut}(T)$ are given by $\{1\},\dots,\{n\}$ (they will sometimes be identified with $1,\dots,n$ to simplify the notation).

We now build the cut-tree $\mathrm{Cut}(T)$ inductively, together with a decoration on the set of its internal nodes. At the r-th step, we let B be the equivalence class for $\sim_{r-1}$ containing the indices i such that $e(r) \in T(i, r-1)$. Deleting the edge $e(r)$ splits this block B in two new equivalence classes B', B'' for $\sim_r$, such that $B = B' \sqcup B''$. We draw two edges between B and the sets B' and B''. We define the decoration at point B by setting $f^-(B) = e(r)^-$, $f^+(B) = e(r)^+$, and $f(B) = \{f^-(B), f^+(B)\}$, where e^- (resp. e^+) denotes the extremity of the edge e which is closest to (resp. furthest from) the root. We stress that, although $f^-(B)$ and $f^+(B)$ depend on the choice of the root, this is not the case for $f(B)$.

Note that with this construction, the graph-distance between the leaf $\{i\}$ and the root in $\mathrm{Cut}(T)$ is the number of cuts happening in the component of T containing v_i until it becomes a single vertex. Moreover, the decoration function f on $C = \mathrm{Cut}(T)$ verifies the following condition:

Condition 1.1. For each internal node b of C, there exists two leaves i, j of C such that b is the most recent common ancestor of i and j, and $f(b) = \{i, j\}$.

It is clear that the cut-tree and the decoration give all the information we need to re-build the initial tree T. Indeed, each two-point set $f(b)$ corresponds to an edge which was erased during the fragmentation procedure. More generally, for any binary tree C, letting $\mathcal{B}(C)$ be the set of its internal nodes, $\mathcal{L}(C)^{[2]}$ be the set of the pairs of leaves of C, and $f : \mathcal{B}(C) \to$ be a decoration function verifying Condition 1.1, we can define a reconstructed *tree* $\mathrm{Rec}(C)$ by setting

$$
\begin{aligned}
V(\mathrm{Rec}(C)) &= \{v_i : i \in \mathcal{L}(C)\} \\
E(\mathrm{Rec}(C)) &= \{\{v_i, v_j\} : \exists b \in \mathcal{B}(C), f(b) = \{i, j\}\}\,.
\end{aligned}
$$

Thus $T = \mathrm{Rec}(\mathrm{Cut}(T))$, with the decoration we described above.

Let us give details on the way we can recover the distance between two vertices in $\mathrm{Rec}(T)$, without building the whole tree. Fix $i, j \in \mathcal{L}(C)$. We build a “subtree” $C^{i,j}$ of C recursively, in the following way:

— Let the root of $C^{i,j}$ be the last common ancestor b of the leaves i and j in C. (If $i = j$, then this is the only vertex of $C^{i,j}$.)

— If $i \neq j$, let i' (resp. j') be the element of $f(b)$ such that b is not on the path between i and i' (resp. j and j'). Build the two subtrees $C^{i,i'}$ and $C^{j,j'}$, and complete the construction by attaching their roots to b.

With this construction, the distance between v_i and v_j in $\mathrm{Rec}(C)$ is the number of internal nodes of $C^{i,j}$. This procedure cannot be applied directly to a continuous tree, but it can be defined “at level ε”: this is what we do in the next section.

Note that in general, the cut-tree will not always be binary. This comes from the fact that in the discrete case, each step consists in cutting at a unique edge (with two ends), whereas in a continuous case, we will cut at points at which several subtrees can be attached. To take this into account in our above model, we should allow deleting several edges at each step (as for example in [28]); but for the reconstruction, this would lead to complications which do not exist in the continuous setting.

1.2 Continuous case

1.2.1 Some definitions on continuous trees

As in the previous chapter, the trees we work with are rooted $\mathbb{R}$-trees. Let us recall some standard definitions and notation.

Let (T, d) be an $\mathbb{R}$-tree. The *multiplicity* $\mathrm{mult}(x)$ of a point $x \in T$ is the number of connected components of $T \setminus \{x\}$ (possibly infinite). We say that x is a leaf of T if $\mathrm{mult}(x) = 1$, and that

x is a branching point of T if $\mathrm{mult}(x) \geq 3$. We write $\mathcal{L}(T)$ for the set of the leaves of T, and $\mathcal{B}(T)$ for the set of its branching points. For all $x, y \in T$, we let $[\![x, y]\!]$ be the unique injective path between x and y, and $]\!]x, y[\![= [\![x, y]\!] \setminus \{x, y\}$. The metric d gives rise to a *length measure* λ on T which is the unique σ-finite measure such that $\lambda([\![u, v]\!]) = d(u, v)$ for all $u, v \in T$.

A rooted $\mathbb{R}$-tree (T, d, ρ) comes with a genealogical order $\prec$ such that $x \prec y$ if and only if $x \in [\![\rho, y]\!]$ and $x \neq y$; in this case, we say that x is an *ancestor* of y. The *most recent common ancestor* $x \wedge y$ of $x, y \in T$ is the point of $\{z : z \prec x, z \prec y\}$ which maximises $d(\rho, z)$. For all $b \in \mathcal{B}(T)$ and $x \succ b$, we let T_b^x be the subtree $\{y \in T : b \notin]\!]x, y[\![\}$ (or, equivalently, the closure of the connected component of $T \setminus \{b\}$ which contains x).

One way to generate a rooted $\mathbb{R}$-tree is to encode it using a continuous *contour function*, as in [33]; this is what we do in the next section for the stable and Brownian random trees. Thus our $\mathbb{R}$-trees are inherently endowed with a mass-measure defined as the pushforward of the Lebesgue measure on $[0, 1]$ under the canonical projection, so they should be seen as metric *measure* spaces with a distinguished point. Moreover, if we limit ourselves to the above setting, two $\mathbb{R}$-trees having "the same shape", but which are not embedded in the same space, are not seen as equal. Therefore, it will be more convenient to define trees as pointed metric measure spaces, seen up to an adequate notion of isometry-equivalence. Therefore, we recall the following definition from Chapter 1.

Definition 1.2. *A pointed metric measure space is a quadruple (X, d, μ, ρ), where μ is a Borel probability measure on the metric space (X, d) and ρ is a point of X. Two such spaces (X, d, μ, ρ), (X', d', μ', ρ') are equivalent if there exists an measurable function $\phi : X \to X'$ such that:*
- *we have $\phi(\rho) = \rho'$.*
- *for all $x, y \in \mathrm{supp}(\mu) \cup \{\rho\}$, we have $d(x, y) = d'(\phi(x), \phi(y))$,*
- *μ' is the pushforward of μ under ϕ,*

The space $\mathbb{M}$ of equivalence classes of complete metric measure spaces (X, d, μ, ρ) such that $X = \mathrm{supp}(\mu) \cup \{\rho\}$, equipped with an adequate topology such as the Gromov–Prokhorov topology, is a Polish space. See for example [36] for details. The equivalence class of a pointed metric measure space (X, d, μ, ρ) is said to be compact if (X, d) is compact.

From now on, we call continuous tree a compact pointed metric measure space (T, d, μ, ρ) such that (T, d) is an $\mathbb{R}$-tree, seen as an element of $\mathbb{M}$. For the random trees we consider, the mass measure μ can be seen as a "uniform" measure on the set of the leaves of T. As a consequence, we say that a random point $l \in T$ is a uniform leaf of T if, conditionally on (T, d, μ, ρ), l is distributed according to μ.

1.2.2 The cut-trees of Brownian and stable trees

Fix $\alpha \in (1, 2]$. Let $X^{(\alpha)}$ be a stable spectrally positive Lévy process with parameter α, such that

$$\mathbb{E}\left[e^{-\lambda X_t^{(\alpha)}}\right] = e^{t\lambda^\alpha} \qquad \forall \lambda \geq 0.$$

For every $t > 0$, let $\hat{X}^{(\alpha,t)}$ be the "reversed" process defined by

$$\hat{X}_s^{(\alpha,t)} = \begin{cases} X_t^{(\alpha)} - X_{(t-s)^-}^{(\alpha)} & \text{if } 0 \leq s < t \\ X_t^{(\alpha)} & \text{if } s = t, \end{cases}$$

and write $\hat{S}_s^{(\alpha,t)} = \sup_{0 \leq r \leq s} \hat{X}_r^{(\alpha,t)}$ for all $r \in [0, t]$. The height process $H^{(\alpha)}$ is the real-valued process such that $H_0^{(\alpha)} = 0$ and, for every $t > 0$, $H_t^{(\alpha)}$ is the local time at level 0 at time t of the process $\hat{X}^{(\alpha,t)} - \hat{S}^{(\alpha,t)}$. The normalization of local time, and the proof of the existence of a continuous modification of this process, are given in [32, Section 1.2].

We say that the *"standard" stable tree* of index α is the $\mathbb{R}$-tree $\mathcal{T}$ coded by the excursion of length 1 of the process $H^{(\alpha)}$, endowed with a probability mass-measure $\mu_{\mathcal{T}}$ which can be seen as the image of the Lebesgue measure on $[0,1]$ under the canonical projection on $\mathcal{T}$. We say that $(\mathcal{T}', d_{\mathcal{T}'}, \mu_{\mathcal{T}'})$ is a stable tree of index α and mass m if the tree $(\mathcal{T}', d_{\mathcal{T}'}/m^{1-1/\alpha}, \mu_{\mathcal{T}'}/m)$ is a standard stable tree of index α. For a such a tree $\mathcal{T}'$, a uniform leaf of $\mathcal{T}'$ is a random point l of $\mathcal{T}'$ such that, conditionally on $\mathcal{T}'$, l is distributed according to $\mu_{\mathcal{T}'}(\cdot)/\mu_{\mathcal{T}'}(\mathcal{T}')$.

Note that the stable tree of index $\alpha = 2$ corresponds to the Brownian tree encoded by $(\sqrt{2}B_t)_{0\leq t\leq 1}$, where B denotes the standard Brownian excursion of length 1. One of the main differences between the Brownian tree and the stable tree of index $\alpha \in (1,2)$ is the fact that the Brownian tree is almost surely binary (i.e. $\mathrm{mult}(b) = 3$ almost surely for all branching point b), whereas a stable tree of index $\alpha \in (1,2)$ only has branching points of infinite multiplicity. This will lead us to adopt slightly different fragmentation mechanisms for these two types of trees. Note that in the latter case, the "size" of a branching point $b \in \mathcal{B}(\mathcal{T})$ can be described by the quantity

$$L(b) = \lim_{\varepsilon\to 0^+} \varepsilon^{-1}\mu\left\{v \in \mathcal{T} : b \in [\![\rho_{\mathcal{T}}, v]\!], d(b,v) < \varepsilon\right\},$$

where $\rho_{\mathcal{T}}$ denotes the root of $\mathcal{T}$. The existence of this quantity was proven in [50] (see also [33]).

From now on, $\mathcal{T}$ will denote a standard stable tree of index α, rooted at a point $\rho_{\mathcal{T}}$. We work with a fixed embedding of the tree $\mathcal{T}$; one can check that the cut-tree we build does not depend on the choice of this embedding. We consider a fragmentation obtained by cutting $\mathcal{T}$ at points and times given by a Poisson process. More precisely:

- If $\alpha = 2$, we let $(t_i, \chi_i)_{i\in\mathbb{N}}$ be the family of the atoms of a Poisson point process with intensity $\Lambda_2(dt, dx) = dt \otimes 2l(dx)$ on $\mathbb{R}_+ \times \mathcal{T}$. (Note that the change in the intensity, compared to [19], comes from the fact that we chose a different normalization.)
- If $\alpha \in (1,2)$, we let $(t_i, \chi_i)_{i\in\mathbb{N}}$ be the family of the atoms of a Poisson point process with intensity $\Lambda_\alpha(dt, dx) = dt \otimes \sum_{b\in\mathcal{B}(\mathcal{T})} L(b)\delta_b(dx)$ on $\mathbb{R}_+ \times \mathcal{T}$.

We say that for all i, at time t_i, the tree $\mathcal{T}$ receives a "cut" at point χ_i. Let $\mathcal{X} = \{\chi_i, i \geq 1\}$.

The cut-tree $\mathrm{Cut}(\mathcal{T})$ describes the genealogy of the connected components created during the fragmentation of $\mathcal{T}$. Informally, we want to establish a correspondence between the points of $\mathcal{T}$ and the leaves of $\mathrm{Cut}(\mathcal{T})$ (other than the root). This can be done in a natural way for the points of $\mathcal{T} \setminus \mathcal{X}$. The points χ_i, $i \in \mathbb{N}$, however, do not belong to the fragmented tree after time t_i, so the corresponding connected component is not well defined. Therefore, for all i, we introduce "copies" $\chi_{i,j}$ of the cut-point χ_i, each of them corresponding to the point χ_i "seen from" one of the connected components of the fragmented tree; thus the points $\chi_{i,j}$, for all j, can be seen as the same point of $\mathcal{T}$, but will be mapped onto different leaves of $\mathrm{Cut}(\mathcal{T})$. Let us explain more precisely how we define these points.

In order to make the Brownian case and the α-stable case fit into the same frame, we let

$$\mathcal{J}_\alpha = \begin{cases} \{0,1\} & \text{if } \alpha = 2 \\ \mathbb{N}\cup\{0\} & \text{otherwise.} \end{cases}$$

(Thus the cardinality of $\mathcal{J}_\alpha$ is equal to the multiplicity of the cut-points in the fragmentation of $\mathcal{T}$.) Fix $i \in \mathbb{N}$. We let $\mathcal{T}_{i,j}$, $j \in \mathcal{J}_\alpha$ denote the connected components of $\mathcal{T} \setminus \{\chi_{i'} : t_{i'} \leq t_i\}$ whose closure (in $\mathcal{T}$) contains χ_i, ordered as follows:

- $\mathcal{T}_{i,0}$ is the component "below" χ_i, i.e. the only component such that for all $x \in \mathcal{T}_{i,0}$, we have $\chi_i \notin [\![\rho_{\mathcal{T}}, x]\!]$,
- the other components $\mathcal{T}_{i,j}$, $j \geq 1$, are ranked by decreasing order of mass.

Formally, we introduce *new* points $\chi_{i,j}$, which correspond to "the point χ_i seen from $\mathcal{T}_{i,j}$", for all $j \in \mathcal{J}_\alpha$. (We could also say that the point χ_i is "split" into the points $\chi_{i,j}$, $j \in \mathcal{J}_\alpha$ at time t_i.) Set

$$\overline{\mathcal{T}}_{\mathcal{X}} = (\mathcal{T} \setminus \mathcal{X}) \sqcup \{\chi_{i,j} : (i,j) \in \mathbb{N} \times \mathcal{J}_\alpha\}.$$

We now come back to the components created during the fragmentation. Our notation should satisfy two criteria: first, we should be able to see each of these components as a compact continuous tree; second, the component containing a point $\chi_{i,j}$, for $(i,j) \in \mathbb{N} \times \mathcal{J}_\alpha$, should be well defined. To this end, these components will be seen as subsets of $\mathcal{T} \sqcup \{\chi_{i,j} : (i,j) \in \mathbb{N} \times \mathcal{J}_\alpha\}$. For all $t \geq 0$ and $x \in \mathcal{T} \setminus \{\chi_i : t_i \leq t\}$, we let $\mathcal{T}(x,t)$ denote the (open) connected component of $\mathcal{T} \setminus \{\chi_i : t_i \leq t\}$ containing x, and we define $\overline{\mathcal{T}}(x,t)$ as the completion of $\mathcal{T}(x,t)$ obtained by adding to $\mathcal{T}(x,t)$ the points $\chi_{i,j}$, for $(i,j) \in \mathbb{N} \times \mathcal{J}_\alpha$ such that:

— $t_i \leq t$,

— χ_i belongs to the closure of $\mathcal{T}(x,t)$,

— $\mathcal{T}(x,t) \subset \mathcal{T}_{i,j}$.

(Note that for a given $i \in \mathbb{N}$, there is at most one of the points $(\chi_{i,j}, j \in \mathcal{J}_\alpha)$ which satisfies these conditions.) We extend this notation naturally to the points $\chi_{i,j}$:

— if $t_i > t$, then for all $j \in \mathcal{J}_\alpha$, we let $\overline{\mathcal{T}}(\chi_{i,j}, t) = \overline{\mathcal{T}}(\chi_i, t)$,

— if $t_i \leq t$, then for all $x \in \mathcal{T}$ such that $\chi_{i,j}$ belongs to $\overline{\mathcal{T}}(x,t)$, we let $\overline{\mathcal{T}}(\chi_{i,j}, t) = \overline{\mathcal{T}}(x,t)$.

Thus, for all $x \in \overline{\mathcal{T}}_{\mathcal{X}}$ and $t \geq 0$, $\overline{\mathcal{T}}(x,t)$ is a (compact) continuous tree. Moreover, one can see that $\overline{\mathcal{T}}(x,t)$ contains a unique point among $\rho_{\mathcal{T}}$ and the $\chi_{i,j}$ for $j \neq 0$. This point will be seen as the root of $\overline{\mathcal{T}}(x,t)$.

Finally, for all $x, y \in \overline{\mathcal{T}}_{\mathcal{X}}$, let $t(x,y) = \inf\left\{t \in \mathbb{R}_+ : \overline{\mathcal{T}}(x,t) \neq \overline{\mathcal{T}}(y,t)\right\}$. If $x, y \in \mathcal{T} \backslash \mathcal{X}$, $t(x,y)$ is the infimum (a.s. finite) of the times when marks appear on the segment $[\![x,y]\!]$; furthermore, if $t(x,y)$ is the *minimum* of such times, we also let $\chi(x,y) = \chi_{i(x,y)}$, with $i(x,y) \in \mathbb{N}$ the index such that $t_{i(x,y)} = t(x,y)$.

We will use the following adaptation of the results of [19, 4] (for the Brownian case) and [28, 2] (for the stable case $\alpha \in (1,2)$):

Proposition 1.3. *There exists a random continuous tree* $\mathrm{Cut}(\mathcal{T})$*, endowed with a distance* $d_{\mathcal{C}}$ *and a probability mass-measure* $\mu_{\mathcal{C}}$*, rooted at a point* $\rho_{\mathcal{C}}$*, which has the following properties:*

1. *Conditionally on* $\mathcal{T}$*, there exists a surjective map* $x \mapsto \ell_x$ *from* $\overline{\mathcal{T}}_{\mathcal{X}}$ *onto* $\mathcal{L}(\mathrm{Cut}(\mathcal{T})) \cup \mathcal{B}(\mathrm{Cut}(\mathcal{T}))$ *such that, for all* $x, y \in \overline{\mathcal{T}}_{\mathcal{X}}$*, we have*
$$d_{\mathcal{C}}(\rho_{\mathcal{C}}, \ell_x) = \int_0^\infty \mu_{\mathcal{T}}(\overline{\mathcal{T}}(x,t))dt,$$
$$d_{\mathcal{C}}(\ell_x, \ell_y) = \int_{t(x,y)}^\infty (\mu_{\mathcal{T}}(\overline{\mathcal{T}}(x,t)) + \mu_{\mathcal{T}}(\overline{\mathcal{T}}(y,t)))dt.$$
2. *The restriction of* ℓ *to* $\mathcal{T} \backslash \mathcal{X}$ *is measurable, and the mass-measure* $\mu_{\mathcal{C}}$ *is the pushforward of* $\mu_{\mathcal{T}}$ *under* ℓ*. Moreover,* $\mu_{\mathcal{C}}$ *is supported on the set of the leaves of* $\mathrm{Cut}(\mathcal{T})$*. In particular, if* $L_0 = \rho_{\mathcal{T}}$ *and* $(L_k, k \geq 1)$ *are independent uniform random leaves of* $\mathcal{T}$*, then* $(\ell_{L_k}, k \geq 0)$ *are independent random leaves of* $\mathrm{Cut}(\mathcal{T})$ *distributed according to* $\mu_{\mathcal{C}}$*.*

Moreover, the tree $(\mathrm{Cut}(\mathcal{T}), d_{\mathcal{C}}, \mu_{\mathcal{C}})$ *has the same distribution as* $(\mathcal{T}, d_{\mathcal{T}}, \mu_{\mathcal{T}})$*.*

Note that this definition of the cut-tree of $\mathcal{T}$ is the same as in [19] for the Brownian case and [28] for the α-stable case (see Section 3.2.3 for details). The only new object we introduce is the "correspondence" ℓ between the points of the initial tree $\mathcal{T}$ and its cut-tree. The main idea of the proof, which will be given in Section 3, consists in iterating the construction of the "first branch" of the cut-tree studied in [4] and [2] (respectively in the case $\alpha = 2$ and in general).

Note that for all $i \in \mathbb{N}$, there exists a unique point $b_i \in \mathcal{B}(\mathrm{Cut}(\mathcal{T}))$ corresponding to the cut-point χ_i: b_i is the most recent common ancestor of the leaves $(\ell_{\chi_{i,j}} : j \in \mathcal{J}_\alpha)$ in $\mathrm{Cut}(\mathcal{T})$.

For any stable tree $\mathcal{T}'$ of index α and mass m, the cut-tree of $\mathcal{T}'$ can be defined exactly as the above proposition, the only difference being that its mass-measure $\mu_{\mathcal{C}'}$ is a finite measure such that $\mu_{\mathcal{C}'}(\mathrm{Cut}(\mathcal{T}')) = m$, instead of a probability measure. In particular, for all $x, y \in \overline{\mathcal{T}}_{\mathcal{X}}$, the subtree $(\mathrm{Cut}(\mathcal{T}))^{\ell_x}_{\ell_x \wedge \ell_y}$ can be identified with the cut-tree of $\overline{\mathcal{T}}(x, t(x,y))$, for the natural fragmentation given by the Poisson point process with atoms $(t_i - t(x,y), \chi_i)$ for $i \geq 1$ such that $t_i > t(x,y)$ and $\chi_i \in \overline{\mathcal{T}}(x, t(x,y))$.

1.2.3 Decoration on $\mathcal{C}$

As in the discrete case, the "reconstructed" tree will be built from a tree $\mathcal{C}$, endowed with a decoration f. Let us first detail what we mean by the term decoration. Note that it is possible to number the branching points $(b_{(i)}, i \in \mathbb{N})$ of $\mathcal{C}$ in a measurable way, for example by decreasing order of width, where the width of a branching point b is the quantity

$$L(b) = \lim_{\varepsilon \to 0^+} \varepsilon^{-1} \mu \{ v \in \mathcal{T} : b \prec v, d(b,v) < \varepsilon \} .$$

(See [33, 50] for the proof of the convergence and details on these quantities.) For all $i \in \mathbb{N}$, we let $(\mathcal{C}^{(i,j)}, j \in \mathcal{J}_\alpha)$ denote the connected components of $\mathcal{C} \setminus \{b_{(i)}\}$, except the one containing the root, ranked by decreasing order of mass. We say that a *decoration* on $\mathcal{C}$ is a sequence of marked points $f = (f_{i,j}, (i,j) \in \mathbb{N} \times \mathcal{J}_\alpha)$ such that, for all $(i,j) \in \mathbb{N} \times \mathcal{J}_\alpha$, we have $f_{i,j} \in \mathcal{C}^{(i,j)}$. For any decoration f, the decorated tree $(\mathcal{C}, f)$ can be seen as a random metric measure space with an countable number of marked points. For our purposes, it will be convenient to see a decoration as a function $f : \mathcal{B}(\mathcal{C}) \to \mathcal{P}(\mathcal{L}(\mathcal{C}))$, where $\mathcal{P}(S)$ denotes the power set of a set S, by letting

$$f(b_{(i)}) = \{ f_{i,j}, j \in \mathcal{J}_\alpha \} .$$

We now explain the choice of the decoration in the case where $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$ for a Brownian tree $\mathcal{T}$, which the closest to what we did in the discrete setting. In this case, we use the function

$$\begin{array}{rccc} f^{\mathcal{T}} : & \mathcal{B}(\mathcal{C}) & \to & \mathcal{P}(\mathcal{L}(\mathcal{C})) \\ & b & \mapsto & \{ f_-^{\mathcal{T}}(b), f_+^{\mathcal{T}}(b) \} \end{array}$$

such that, for all $i \in \mathbb{N}$,

$$f_-^{\mathcal{T}}(b_i) = \ell_{\chi_{i,0}} \quad \text{and} \quad f_+^{\mathcal{T}}(b_i) = \ell_{\chi_{i,1}} .$$

Thus, for all $x, y \in \mathcal{T} \setminus \mathcal{X}$ such that $\ell_x, \ell_y \in \mathcal{L}(\mathcal{C})$, we have that $f^{\mathcal{T}}(\ell_x \wedge \ell_y)$ is the set of the leaves of $\mathrm{Cut}(\mathcal{T})$ corresponding the points into which the cut-point $\chi(x,y)$ is split. (Note that it will be clear from the proof of proposition 1.3 that under the condition $\ell_x, \ell_y \in \mathcal{L}(\mathcal{C})$, the point $\chi(x,y)$ is well-defined.) More generally, if $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$ for a stable tree $\mathcal{T}$ (of index $\alpha \in (1,2]$), the decoration is the function $f^{\mathcal{T}} : \mathcal{B}(\mathcal{C}) \to \mathcal{P}(\mathcal{L}(\mathcal{C}))$ such that, for all $i \in \mathbb{N}$,

$$f^{\mathcal{T}}(b_i) = \{ \ell_{\chi_{i,j}} : (i,j) \in \mathbb{N} \times \mathcal{J}_\alpha \} .$$

Note that this decoration function remains the same if we reroot $\mathcal{T}$ at a uniform leaf.

In the general case where $\mathcal{C}$ is a stable tree of index $\alpha \in (1,2]$, we will use a decoration whose distribution is characterized by the following condition.

Condition 1.4. Conditionally on $\mathcal{C}$:
- The marks $(f_{i,j}, (i,j) \in \mathbb{N} \times \mathcal{J}_\alpha)$ are mutually independent.
- For all $(i,j) \in \mathbb{N} \times \mathcal{J}_\alpha$, $f_{i,j}$ is a uniform leaf of $\mathcal{C}^{(i,j)}$.

Note that the order in which the branching points and the connected components $\mathcal{C}^{(i,j)}$ are numbered do not matter to check this condition. If we see f as a function from $\mathcal{B}(\mathcal{C})$ into $\mathcal{P}(\mathcal{L}(\mathcal{C}))$, this condition has the following equivalent formulation.

Condition 1.5. For all $b \in \mathcal{B}(\mathcal{C})$, we write $(\mathcal{C}_b^{(j)}, j \in \mathcal{J}_\alpha)$ for the connected components of $\mathcal{C} \setminus \{b\}$ which do not contain the root. Conditionally on $\mathcal{C}$, for all $b \in \mathcal{B}(\mathcal{C})$, we have $f(b) = \{ f^{(j)}(b), j \in \mathcal{J}_\alpha \}$, where:
- The random variables $f^{(j)}(b)$, $j \in \mathcal{J}_\alpha$ are mutually independent.
- For all $j \in \mathbb{N}$, $f^{(j)}(b)$ is a uniform leaf of $\mathcal{C}_b^{(j)}$.

Moreover, the random sets $f(b)$, $b \in \mathcal{B}(\mathcal{C})$ are mutually independent.

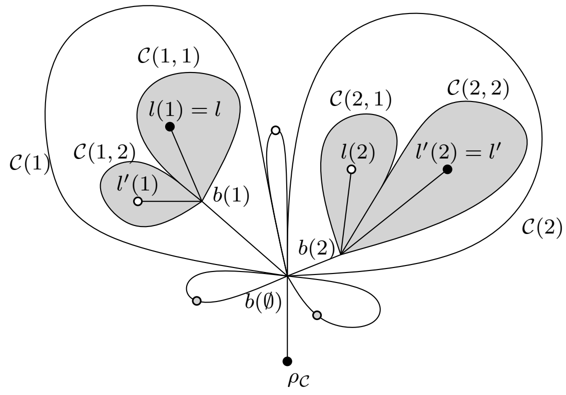

Figure 3.1 – The subtrees $(\mathcal{C}(u), u \in \{1,2\}^2)$ for two given leaves l and l'. The points of $f(b(\emptyset))$ are represented by white vertices.

In particular, for all $b \in \mathcal{B}(\mathcal{C})$ and $l \succ b$, we call $f_l(b)$ the unique element of $f(b) \cap \mathcal{C}_b^l$. Note that if l and l' are independent uniform leaves of $\mathcal{C}$, letting $b = l \wedge l'$, then conditionally on $\mathcal{C}_b^l$ and $\mathcal{C}_b^{l'}$, $f_l(b)$ and $f_{l'}(b)$ are independent uniform leaves of $\mathcal{C}_b^l$ and $\mathcal{C}_b^{l'}$ respectively, both independent of l, l'.

The fact that the above decoration function $f^{\mathcal{T}}$ verifies the above condition will be shown in Section 4.1.

1.2.4 Construction of $\mathcal{C}^{l,l'}(\varepsilon)$

For all $l, l' \in \mathcal{L}(\mathcal{C})$ and $\varepsilon > 0$, we build a subtree of $\mathcal{C}$ which will encode the necessary information to compute a "reconstructed distance" between l and l', "at level ε". The distances in the reconstructed tree will be obtained as scaling limits of these quantities as $\varepsilon \to 0$.

We first introduce some notation. Let $U = \{\emptyset\} \cup \bigcup_{n\in\mathbb{N}} \{1,2\}^n$. For all $u, u' \in U$, we write (u, u') for the concatenation of the sequences u and u'. We will see discrete binary trees as subsets of U, by interpreting $\emptyset$ as the root-vertex, and $(u,1), (u,2)$ as the children of an individual u. If $u = (u_1, \ldots, u_n)$, with $n \geq 1$, let $p(u)$ denote the parent of the individual u: $p(u) = (u_1, \ldots, u_{n-1})$ if $n \geq 2$, and $p(u) = \emptyset$ if $n = 1$. Such trees have a natural planar structure, but it will not have any importance in our construction.

Fix $l, l' \in \mathcal{L}(\mathcal{C}) \setminus \{\rho_{\mathcal{C}}\}$. We build recursive sequences $l(u)$, $l'(u)$, $b(u)$ and $\mathcal{C}(u)$, indexed by the elements of U, in the following way:

— Let $l(\emptyset) = l$, $l'(\emptyset) = l'$, and $\mathcal{C}(\emptyset) = \mathcal{C}$.

— For all $u \in U$, let $b(u) = l(u) \wedge l'(u)$, and

$$\begin{aligned} l(u,1) &= l(u) \qquad & l'(u,1) &= f_{l(u)}(b(u)) \qquad & \mathcal{C}(u,1) &= (\mathcal{C}(u))_{b(u)}^{l(u)} \\ l(u,2) &= f_{l'(u)}(b(u)) \qquad & l'(u,2) &= l'(u) \qquad & \mathcal{C}(u,2) &= (\mathcal{C}(u))_{b(u)}^{l'(u)}. \end{aligned}$$

We see $\mathcal{C}(u,1)$ and $\mathcal{C}(u,2)$ as trees rooted at $b(u)$.
The first steps of this construction are illustrated on Figure 3.1.

Fix $\varepsilon > 0$. Let $U_\varepsilon = \{u \in U : \mu_{\mathcal{C}}(\mathcal{C}(u)) \geq \varepsilon\}$. We define $\mathcal{C}^{l,l'}(\varepsilon)$ as the subtree of $\mathcal{C}$ generated by the root and the leaves $l(u)$, $l'(u)$, $u \in U_\varepsilon$:

$$\mathcal{C}^{l,l'}(\varepsilon) = \bigcup_{u \in U_\varepsilon} \left([\![\rho_{\mathcal{C}}, l(u)]\!] \cup [\![\rho_{\mathcal{C}}, l'(u)]\!] \right).$$

Thus, $\mathcal{C}^{l,l'}(\varepsilon)$ is an $\mathbb{R}$-tree whose branching points of $\mathcal{C}^{l,l'}(\varepsilon)$ are exactly the $b(u)$, $u \in U_\varepsilon$. We will mainly use the discrete structure of this tree, seeing $\mathcal{C}^{l,l'}(\varepsilon)$ as the *discrete* binary tree whose internal vertices are the $u \in U_\varepsilon$. The quantity we are interested in is the number of branching points of $\mathcal{C}^{l,l'}(\varepsilon)$, denoted by $N^{l,l'}(\varepsilon)$. Note that with the terminology of discrete trees, this corresponds to the number of internal nodes, as in the discrete case.

In the case where $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$, let us explain informally why $N^{\ell_x,\ell_y}(\varepsilon)$ is linked with the distance $d_{\mathcal{T}}(x,y)$, for any two points $x, y \in \mathcal{T} \setminus \mathcal{X}$. Consider the fragmentation "at level ε" of $\mathcal{T}$, obtained by keeping only the cut-points χ_i such that $\mu_{\mathcal{T}}(\overline{\mathcal{T}}(\chi_i, t_i^-)) \geq \varepsilon$. In other words, we stop the fragmentation in a component as soon as its mass becomes lower than ε. Then $N^{\ell_x,\ell_y}(\varepsilon)$ corresponds to the number of cuts that happen on the path $[\![x,y]\!]$, for the fragmentation at level ε, so it is natural that this should approximate the length of this path.

1.3 Main results

From now on, we assume that $\mathcal{C}$ is a stable tree of index $\alpha \in (1,2]$, endowed with a decoration f which verifies Condition 1.4. The first step of the reconstruction is to prove that for two independent uniform leaves l, l' of $\mathcal{C}$, the rescaled quantities $N^{l,l'}(\varepsilon)$ actually converge to a quantity which will be our new distance.

Proposition 1.6. *Let l, l' be independent uniform leaves of $\mathcal{C}$. There exists a random variable $\delta_{\mathcal{C}}(l,l')$ such that we have the almost sure convergence*

$$\frac{(\alpha-1)^2 c_\alpha}{2} \cdot \varepsilon^{1-1/\alpha} N^{l,l'}(\varepsilon) \xrightarrow[\varepsilon \to 0]{} \delta_{\mathcal{C}}(l,l'),$$

where $c_\alpha = \Gamma(1-1/\alpha)/(\alpha\Gamma(2-2/\alpha))$. Moreover, $\delta_{\mathcal{C}}(l,l')$ and $d_{\mathcal{C}}(l,l')$ have the same distribution.

Note that the last fact will not be used afterwards, since it will be a consequence of the following proposition. Nevertheless, it can also be shown directly, as will be done in Section 2.

In the case where $\mathcal{C}$ is the cut-tree of a stable tree, the reconstructed distance exactly corresponds to the distance in the initial tree:

Proposition 1.7. *If $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$, where $\mathcal{T}$ is a stable tree of index α, and if L, L' are independent uniform leaves of $\mathcal{T}$, then we have*

$$\delta_{\mathrm{Cut}(\mathcal{T})}(\ell_L, \ell_{L'}) = d_{\mathcal{T}}(L, L') \qquad a.s. \tag{3.1}$$

This special case is particularly important, since the identity (3.1) will allow us to characterise the joint distribution of the reconstructed distances between i.i.d. points of $\mathcal{C}$. Indeed, letting $\mathcal{T}$ be a stable tree of index α, we have the equality in distribution

$$(\mathcal{C}, f) \overset{(d)}{=} (\mathrm{Cut}(\mathcal{T}), f^{\mathcal{T}}).$$

Note that if we add a distinguished leaf l^0 to the tree $\mathcal{C}$, uniform and independent of f, this gives

$$(\mathcal{C}, f, l^0) \overset{(d)}{=} (\mathrm{Cut}(\mathcal{T}), f^{\mathcal{T}}, \ell_{\rho_{\mathcal{T}}}).$$

The distinguished leaf l^0 will correspond to the root of the reconstructed tree. Now set $l_0 = l^0$, and let $(l_k)_{k\geq 1}$ be an i.i.d. sequence with law $\mu_{\mathcal{C}}$, independent of the decoration and of l^0, conditionally given $\mathcal{C}$. Similarly, set $L_0 = \rho_{\mathcal{T}}$ and let $(L_k)_{k\geq 1}$ be an i.i.d. sequence with law $\mu_{\mathcal{T}}$, independent of the decoration and of $\rho_{\mathcal{T}}$, conditionally given $\mathcal{T}$. For all $\varepsilon > 0$, we have

$$(N^{l_i,l_j}(\varepsilon))_{i,j\geq 0} \overset{(d)}{=} (N^{\ell_{L_i},\ell_{L_j}}(\varepsilon))_{i,j\geq 0},$$

hence

$$(\delta_{\mathcal{C}}(l_i,l_j))_{i,j\geq 0} \overset{(d)}{=} (\delta_{\mathrm{Cut}(\mathcal{T})}(\ell_{L_i},\ell_{L_j}))_{i,j\geq 0}.$$

Together with (3.1), this yields

$$(\delta_{\mathcal{C}}(l_i, l_j))_{i,j\geq 0} \overset{(d)}{=} (d_{\mathcal{T}}(L_i, L_j))_{i,j\geq 0}. \tag{3.2}$$

As a consequence, conditionally on $(\mathcal{C}, f)$, the function δ is a distance on $\{l_k, k \geq 0\}$ and, using the terminology of Aldous [7], the family of the reduced trees $\mathcal{R}(k) := (\{l_j, 0 \leq j \leq k\}, \delta)$ forms a *consistent* family of random rooted trees which satisfies the *leaf-tight condition*:

$$\min_{1\leq j\leq k} \delta(l_0, l_j) \xrightarrow[k\to\infty]{\mathbb{P}} 0.$$

This implies that there exists a random continuous tree $\mathrm{Rec}(\mathcal{C})$ for which, conditionally on $(\mathcal{C}, f)$, the distribution of the reduced tree formed by the root and the first k leaves of an i.i.d. sequence, for $k \in \mathbb{N}$, is the same as the distribution of $(\mathcal{R}(k), k \in \mathbb{N})$. This tree is called the reconstructed tree of $\mathcal{C}$.

Note that $\mathrm{Rec}(\mathcal{C})$ depends on $\mathcal{C}$ and on the extra randomness of the decoration, in the same way as $\mathrm{Cut}(\mathcal{T})$ depends on $\mathcal{T}$ and on the extra randomness of the Poisson point process. Besides, the role played by the root of the initial tree is not the same in these two operations: the decorated cut-tree $(\mathrm{Cut}(\mathcal{T}), f^{\mathcal{T}})$ does not depend on the root of $\mathcal{T}$, and remains a rooted tree even for a non-rooted initial tree, whereas the reconstructed tree $\mathrm{Rec}(\mathcal{C})$ can only be built for a rooted tree $\mathcal{C}$ (in particular, the condition 1.4 under which f is a suitable decoration depends on the genealogical order on $\mathcal{C}$). The distinguished leaf l^0 we added above to perform the reconstruction only serves to give us a rooting convention on $\mathrm{Rec}(\mathcal{C})$, and to show that we can recover all the information on the initial tree if it is rooted.

We can now state our main result:

Theorem 1.8. *The reconstructed tree* $\mathrm{Rec}(\mathcal{C})$ *is a standard stable tree of index* α*. Moreover, if* $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$ *for a stable tree* $\mathcal{T}$ *of index* α*, then* $\mathrm{Rec}(\mathrm{Cut}(\mathcal{T})) = \mathcal{T}$ *almost surely.*

This result is a straightforward consequence of the identity (3.2) and of the construction of $\mathrm{Rec}(\mathcal{C})$. Note that without using the fact that $\mathcal{C}$ has the same distribution as the cut-tree of a stable tree $\mathcal{T}$ of index α, the proof of this theorem would be much more involved. We would first have to use the construction of the $N_\varepsilon^{l,l'}$ (instead of the identity (3.2)) to show that we can build a tree $\mathrm{Rec}(\mathcal{C})$ which has the above matrix of mutual distances. Then, we would need to get a characterization of the joint distribution of the $\delta_{\mathcal{C}}(l_i, l_j)$, $i, j \geq 0$, for example by using recursive distributional equations as we will do in the two-leaf case, in Section 2.

To complement this result, we will finally show that we can define mappings $r_{\mathcal{C}} : \mathcal{B}(\mathcal{C}) \to \mathrm{Rec}(\mathcal{C})$ and $t_{\mathcal{C}} : \mathcal{B}(\mathcal{C}) \to \mathbb{R}_+$, depending only on $(\mathcal{C}, f)$, such that the following property holds:

Theorem 1.9. *The points* $(r_{\mathcal{C}}(b), t_{\mathcal{C}}(b))$*,* $b \in \mathcal{B}(\mathcal{C})$*, have the same distribution as the atoms of a Poisson point process with intensity* $\Lambda_\alpha(dt, dx)$ *on* $\mathbb{R}_+ \times \mathrm{Rec}(\mathcal{C})$*. Moreover, for the fragmentation induced by these atoms, we have* $\mathrm{Cut}(\mathrm{Rec}(\mathcal{C})) = \mathcal{C}$ *almost surely.*

The rest of the article is organised as follows. In Section 2, we focus on the reconstructed distance between two uniform leaves, giving the proof of Proposition 1.6. Section 3 gives the necessary details on the construction of $\mathrm{Cut}(\mathcal{T})$ we announced in Proposition 1.3. Section 4 is devoted to the case where $\mathcal{C}$ is the cut-tree of a stable tree, and in particular to the proof of Proposition 1.7. Finally, in Section 5, we give a correspondence between the leaves of $\mathcal{C}$ (except the root) and a dense subset of points of $\mathrm{Rec}(\mathcal{C})$; in particular, we use this correspondence to give the proof of Theorem 1.9 (a more detailed version of this result is given in Proposition 5.9.)

From now on, except in section 5, we will omit the index in the notation $\delta_{\mathcal{C}}$ (we always consider the reconstructed distance of $\mathcal{C}$, and sometimes study the case where $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$).

2 Reconstructed distance between two uniform leaves

2.1 Distributions of masses and distances in a stable tree

2.1.1 Distance between two uniform points

Let us first recall the definition of the Mittag-Leffler distribution. For $\beta \in (0,1)$, let σ_β be a stable random variable having Laplace transform

$$\mathbb{E}\left[\exp(-\lambda\sigma_\beta)\right] = \exp(-\lambda^\beta), \quad \lambda \geq 0.$$

For $\theta > -\beta$, we say that a random variable M has the *generalized Mittag–Leffler distribution* $\mathrm{ML}(\beta,\theta)$ if, for all suitable test functions φ,

$$\mathbb{E}\left[f(M)\right] = C_{\beta,\theta}\mathbb{E}\left[\sigma_\beta^{-\theta}\varphi\left(\sigma_\beta^{-\beta}\right)\right],$$

where $C_{\beta,\theta}$ is the appropriate normalizing constant. (See Goldschmidt and Haas [35] for much more about this distribution and its relationship to the stable trees.)

It is known from [32, Theorem 3.3.3] that the distance between two uniform points of a stable tree of index α and total mass m is distributed as $m^{1-1/\alpha}(1/\alpha \cdot \Delta)$, with Δ a random variable following the Mittag–Leffler distribution $\mathrm{ML}(1-1/\alpha, 1-1/\alpha)$. When $\alpha = 2$, the law of the distance between two uniform points was earlier proved to follow the Rayleigh distribution, by Aldous [7]. This is in agreement with the current result, up to a choice of normalization.

2.1.2 Cutting a stable tree at a branching point between two uniform leaves

For all $n \in \mathbb{N}$ and $\theta_1, \dots, \theta_n > 0$, write $\mathcal{H}_n = \left\{(x_1,\dots,x_n) \in \mathbb{R}_+^n : \sum_{i=1}^n x_i = 1\right\}$. A $\mathcal{H}_n$-valued random vector $(X_1,\dots,X_n)$ has the Dirichlet distribution $\mathrm{Dir}(\theta_1,\dots,\theta_n)$ if its density with respect to Lebesgue measure on $\mathcal{H}_n$ is

$$\frac{\Gamma(\sum_{i=1}^n \theta_i)}{\prod_{i=1}^n \Gamma(\theta_i)} \prod_{j=1}^n x_i^{\theta_i - 1}.$$

Let $(\mathcal{S}, d, \mu, \rho)$ be a stable tree of index $\alpha \in (1,2]$ conditioned to have total mass $m > 0$. Let x, y be independent uniform leaves of $\mathcal{S}$, and $b = x \wedge y$. Recall that $\mathcal{S}_b^x$, for example, denotes the connected component of $\mathcal{S} \setminus \{b\}$ which contains x. The trees $\mathcal{S}_b^x$ and $\mathcal{S}_b^y$ are naturally endowed with the restrictions of the distance d and mass-measure μ.

It is known from [8] and [38] (see in particular Corollary 10 in the latter) that:

Theorem 2.1. *1. The mass ratios $\mu(\mathcal{S}_b^x)/m$, $\mu(\mathcal{S}_b^y)/m$, $(m - \mu(\mathcal{S}_b^x) - \mu(\mathcal{S}_b^y))/m$ follow a Dirichlet distribution* $\mathrm{Dir}(1-1/\alpha, 1-1/\alpha, 1/\alpha)$.

2. *Conditionally on these masses, $(\mathcal{S}_b^x, b, x)$ and $(\mathcal{S}_b^y, b, y)$ are distributed as independent stable trees of index α with two independent leaves.*
3. *Conditionally on their masses, the closures of the other connected components of $\mathcal{S} \setminus \{b\}$ above b are independent stable trees of index α, independent of $\mathcal{S}_b^x$ and $\mathcal{S}_b^y$.*

2.1.3 A recursive distributional equation

Writing $d(x,y) = d_{\mathcal{S}_b^x}(x,b) + d_{\mathcal{S}_b^y}(b,y)$, Theorem 2.1 yields the distributional equation

$$X_1^{1-1/\alpha}\Delta^{(1)} + X_2^{1-1/\alpha}\Delta^{(2)} \overset{(d)}{=} \Delta$$

where $(X_1, X_2, 1 - X_1 - X_2)$ follows the Dirichlet distribution $\mathrm{Dir}(1-1/\alpha, 1-1/\alpha, 1/\alpha)$, and $\Delta^{(1)}, \Delta^{(2)}$ are two independent Mittag–Leffler $\mathrm{ML}(1-1/\alpha, 1-1/\alpha)$ variables. Note that this also follows from Proposition 4.2 of Goldschmidt and Haas [35], or from a simple moment calculation (the generalized Mittag–Leffler distributions being characterized by their moments). We now show that this equation characterizes the Mittag–Leffler distribution $\mathrm{ML}(1-1/\alpha, 1-1/\alpha)$:

Lemma 2.2. *Suppose that* $(X_1, X_2, X_3) \sim \mathrm{Dir}(1-1/\alpha, 1-1/\alpha, 1/\alpha)$ *and that* $\Delta^{(1)}, \Delta^{(2)}$ *are independent copies of some* $\mathbb{R}_+$*-valued random variable* Δ*. Then* $\Delta \sim \mathrm{ML}(1-1/\alpha, 1-1/\alpha)$ *is a solution of the recursive distributional equation*

$$X_1^{1-1/\alpha}\Delta^{(1)} + X_2^{1-1/\alpha}\Delta^{(2)} \overset{(d)}{=} \Delta, \tag{3.3}$$

and this solution is unique up to multiplication by a non-negative constant.

Proof. The left-hand side of (3.3) is an instance of the *smoothing transform* applied to the law of Δ. The fixed points of the smoothing transform have been completely characterized by Durrett and Liggett [34]. Indeed, the space of fixed points is determined by the analytical properties of the function $F : \mathbb{R}_+ \to \mathbb{R}$ defined (in our setting) by

$$F(s) = \log\left(\mathbb{E}\left[X_1^{(1-1/\alpha)s}\mathbf{1}_{X_1>0} + X_2^{(1-1/\alpha)s}\mathbf{1}_{X_2>0}\right]\right),$$

for $s \geq 0$. Since X_1 and X_2 are marginally both distributed as $\mathrm{Beta}(1-1/\alpha, 1)$, it is easily checked that $X_1^{1-1/\alpha} \overset{(d)}{=} X_2^{1-1/\alpha} \overset{(d)}{=} U$, where U is uniform on $[0,1]$. Hence,

$$F(s) = \log\left(2\mathbb{E}\left[U^s\right]\right) = \log(2) - \log(s+1), \quad s \geq 0,$$

for any $\alpha \in (1,2]$. Observe that F has its unique zero in $(0,1]$ at $s=1$, and that $F'(1) = -1/2 < 0$. In this case, part (a) of Theorem 2 of [34] entails that (3.3) has a unique distributional solution, up to multiplication by a non-negative constant. □

2.2 Reconstructed distance between two uniform leaves

From now on, we let l, l' be two independent uniform leaves of $\mathcal{C}$. The goal of this section is to show the almost-sure convergence of $\varepsilon^{1-1/\alpha} \cdot N^{l,l'}(\varepsilon)$. Recalling the notation of section 1.2.4, we build a fragmentation with a discrete genealogical structure given by the tree $\mathcal{C}^{l,l'}(\varepsilon)$. More precisely, we mark each individual $u \in U$ with the mass $m(u) = \mu_{\mathcal{C}}(\mathcal{C}(u))$. This represents a fragmentation in which each fragment u has mass $m(u)$, and breaks into two fragments $(u,1)$ and $(u,2)$. Let us first see that this fragmentation is homogeneous.

2.2.1 Characteristics of the fragmentation

We show recursively that $\mathcal{C}(u)$ is a stable tree of index α and mass $m(u)$, with two independent uniform leaves $l(u)$, $l'(u)$. This is true for $u = \emptyset$. Fix $u \in U$, and assume this property holds for u. We have $\mathcal{C}(u,1) = (\mathcal{C}(u))_{l(u)\wedge l'(u)}^{l(u)}$ and $\mathcal{C}(u,2) = (\mathcal{C}(u))_{l(u)\wedge l'(u)}^{l'(u)}$. Thus these trees are obtained by cutting $\mathcal{C}(u)$ into three parts as in Theorem 2.1. It follows from the proposition that:

- The mass ratios $m(u,1)/m(u)$, $m(u,2)/m(u)$, $1-(m(u,1)+m(u,2))/m(u)$ follow a Dirichlet distribution $\mathrm{Dir}(1-1/\alpha, 1-1/\alpha, 1/\alpha)$.
- Conditionally on these masses, $\mathcal{C}(u,1)$ is a stable tree of index α with a uniform leaf $l(u)$. Since $l'(u,1) = f_{l(u)}(b(u))$, $l(u,1)$ and $l'(u,1)$ are independent uniform leaves of $\mathcal{C}(u,1)$. Similarly, $\mathcal{C}(u,2)$ is a stable tree of index α and mass $m(u,2)$, with two independent uniform leaves $l(u,2), l'(u,2)$, independent of $\mathcal{C}(u,1)$.

Moreover, the decorations on $\mathcal{C}(u,1)$ and $\mathcal{C}(u,2)$ are independent, and independent of $f(b(u))$. As a consequence, the different fragments of the same generation evolve independently, and for all $u \in U$, for any continuous function g on $[0,1]$, we have

$$\mathbb{E}\left[g\left(\frac{m(u,1)}{m(u)}, \frac{m(u,2)}{m(u)}\right)\right] = \int_{[0,1]^2} g(s_1,s_2)\nu_\alpha(ds_1,ds_2),$$

where

$$\nu_\alpha(ds_1,ds_2) = \frac{1-1/\alpha}{\Gamma(1/\alpha)\Gamma(1-1/\alpha)}\frac{\mathbf{1}_{s_1+s_2<1}ds_1ds_2}{s_1^{1/\alpha}s_2^{1/\alpha}(1-s_1-s_2)^{1-1/\alpha}}.$$

This shows that the fragmentation is homogeneous, with dislocation law ν_α.

2.2.2 Almost-sure convergence of $\varepsilon^{1-1/\alpha} \cdot N^{l,l'}(\varepsilon)$

We now use the results of Bertoin and Martinez [18] on fragmentation energies, and of Nerman [53], to get the convergence of a rescaled version of $N^{l,l'}(\varepsilon)$. Adopting the setting of [18], we assume that each dislocation of a fragment of mass m into smaller fragments of masses m_1, m_2, has an energy cost of the form $m^\beta \varphi(m_1/m, m_2/m)$. The total energy cost of the process, if we "freeze" each fragment as soon as it reaches a mass lower than ε, can be written as

$$E(\varepsilon) = \sum_{u \in U} \mathbf{1}_{m(u) \geq \varepsilon} m(u)^\beta \varphi\left(\frac{m(u,1)}{m(u)}, \frac{m(u,2)}{m(u)}\right),$$

In particular, our quantity

$$N^{l,l'}(\varepsilon) = \sum_{u \in U} \mathbf{1}_{m(u) \geq \varepsilon}$$

can be seen as such a fragmentation energy, with $\varphi(s_1, s_2) = 1$ for all s_1, s_2, and $\beta = 0$. Note that although β was supposed to be positive in [18], it is easy to check that the results they obtained in the discrete framework also hold for $\beta = 0$. Simple calculations show that, for $\alpha^* = 1 - 1/\alpha$,

$$\int_{[0,1]^2} (s_1^{\alpha^*} + s_2^{\alpha^*}) \nu_\alpha(ds_1, ds_2) = 1,$$

and

$$\int_{[0,1]^2} (s_1^{\alpha^*} + s_2^{\alpha^*}) \log((s_1^{\alpha^*} + s_2^{\alpha^*}) \wedge 1) \nu_\alpha(ds_1, ds_2) < \infty,$$

which means that the fragmentation satisfies the Malthusian hypothesis with parameter α^*. Since

$$\int_{[0,1]^2} |\varphi(s_1, s_2)| \, \nu_\alpha(ds_1, ds_2) = 1 < \infty,$$

we can apply Theorem 1 of [18]:

$$\varepsilon^{\alpha^*} N^{l,l'}(\varepsilon) \underset{\varepsilon \to 0}{\longrightarrow} \frac{W_\infty}{\alpha^* c(\alpha^*)} \int_{[0,1]^2} |\varphi(s_1, s_2)| \, \nu_\alpha(ds_1, ds_2) \qquad \text{in } L^1,$$

where

$$c(\alpha^*) = \int_{[0,1]^2} \left(s_1^{\alpha^*} \log\left(\frac{1}{s_1}\right) + s_2^{\alpha^*} \log\left(\frac{1}{s_2}\right)\right) \nu_\alpha(ds_1, ds_2),$$

and W_∞ is the almost sure limit of the martingale

$$W_n = \sum_{u \in \{1,2\}^n} m(u)^{\alpha^*} = \sum_{u \in \{1,2\}^n} m(u)^{1-1/\alpha}.$$

Calculations show that $c(\alpha^*) = \alpha(\alpha - 1)/2$. As a consequence, we have the convergence

$$\varepsilon^{1-1/\alpha} N^{l,l'}(\varepsilon) \underset{\varepsilon \to 0}{\longrightarrow} \frac{2}{(\alpha-1)^2} W_\infty \tag{3.4}$$

in L^1. We now apply further results of Nerman to show that this convergence holds almost surely.

To this end, note that we can also write

$$N^{l,l'}(\varepsilon) = \sum_{u \in U} \phi(t - \sigma_u),$$

where $t = -\log \varepsilon$, $\sigma_u = -\log m(u)$ and $\phi = \mathbf{1}_{\mathbb{R}_+}$, so that $N_\varepsilon^{l,l'}$ corresponds to the quantity Z_t^ϕ studied in [53]. Theorem 5.4 of [53] now gives the almost sure convergence of this quantity under two conditions:

1. the existence of an integrable, nonincreasing positive function h on $\mathbb{R}_+$ such that

$$\sup_t \frac{e^{-\alpha^* t}\phi(t)}{h(t)}$$

has finite expectation. In our case ϕ is deterministic, so it suffices to find h such that this quantity is finite. For example $h(t) = e^{-\alpha^* t}$ works.

2. a similar condition on a quantity depending on the dislocation law, which is seen to be weaker than

$$\int_{[0,1]^2} \left((\log s_1)^2 s_1^{\alpha^*} + (\log s_2)^2 s_2^{\alpha^*} \right) \nu_\alpha(ds_1, ds_2) < \infty.$$

(see [53, Equation (5.7)]). For our dislocation measure, this quantity is equal to

$$\left(1 - \frac{1}{\alpha}\right) \int_0^1 \frac{(\log s)^2}{s^{2/\alpha - 1}} ds < \infty.$$

2.2.3 Distribution of the limit and conclusion

Since the above fragmentation is homogeneous, for all $n \in \mathbb{N}$, W_n has the same law as

$$m(1)^{1-1/\alpha} W_{n-1}^{(1)} + m(2)^{1-1/\alpha} W_{n-1}^{(2)},$$

where $W_{n-1}^{(1)}$ and $W_{n-1}^{(2)}$ are two independent copies of W_{n-1}. Taking $n \to \infty$, this shows that W_∞ is a solution of the distributional equation (3.3). Therefore, Lemma 2.2 implies the existence of a constant c such that $W_\infty = c\Delta$, with $\Delta \sim \mathrm{ML}(1 - 1/\alpha, 1 - 1/\alpha)$. We have $\mathbb{E}[W_\infty] = 1$ and

$$\mathbb{E}[\Delta] = \frac{\Gamma(1 - 1/\alpha)}{\Gamma(2 - 2/\alpha)},$$

so

$$W_\infty \overset{(d)}{=} \frac{\Gamma(2 - 2/\alpha)}{\Gamma(1 - 1/\alpha)} \Delta \overset{(d)}{=} \frac{1}{c_\alpha} d_{\mathcal{C}}(l, l').$$

Since the almost sure convergence (3.4) is equivalent to

$$\frac{(\alpha - 1)^2 c_\alpha}{2} \cdot \varepsilon^{1-1/\alpha} N^{l,l'}(\varepsilon) \xrightarrow[\varepsilon \to 0]{} \delta_{\mathcal{C}}(l, l') := c_\alpha \cdot W_\infty \qquad \text{a.s.},$$

this completes the proof of Proposition 1.6.

3 Construction of $\mathrm{Cut}(\mathcal{T})$

3.1 Main idea

As mentioned in the Introduction, the main idea of the overall construction consists in iterating the construction of the “first branch” of the cut-tree studied in [4, 2]. In the latter construction, one only keeps the cut-points which fall into the component containing the root of $\mathcal{T}$; the components which are above such cut-points thus play a different role than the ones below. To iterate this construction, we use all the cut-points, but this asymmetry still appears.

As a consequence, we let

$$\mathcal{J}_\alpha^+ = \mathcal{J}_\alpha \setminus \{0\} = \begin{cases} \{1\} & \text{if } \alpha = 2 \\ \mathbb{N} & \text{otherwise,} \end{cases}$$

so that the components above a cut-point χ_i are the components containing the points $\chi_{i,j}$ for $j \in \mathcal{J}_\alpha^+$. For all $(i,j) \in \mathbb{N} \times \mathcal{J}_\alpha^+$, let $\overline{\mathcal{T}}_{i,j}$ be the tree $\overline{\mathcal{T}}(\chi_{i,j}, t_i)$, $r_{i,j} = \chi_{i,j}$ be its root and $t_{i,j} = t_i$. We also set $\overline{\mathcal{T}}_{0,0} = \mathcal{T}$, $r_{0,0} = \rho_\mathcal{T}$ and $t_{0,0} = 0$. We let $\mathcal{E}_\alpha = \{(0,0)\} \cup (\mathbb{N} \times \mathcal{J}_\alpha^+)$ be the set of the indices of these components. For all $i \in \mathbb{N}$, there exists a unique element $p(i)$ of $\mathcal{E}_\alpha$ such that

$$t_i > t_{p(i)} \quad \text{and} \quad \chi_i \in \overline{\mathcal{T}}(r_{p(i)}, t_i^-).$$

We think of $p(i)$ as the "parent" of i, since the cut which creates the subtrees $\overline{\mathcal{T}}_{i,j}$, $j \in \mathcal{J}_\alpha^+$ happens in the component of $r_{p(i)}$. Finally, for all $(i,j) \in \mathcal{E}_\alpha$, let

$$\delta_{(i,j)} = \int_{t_i}^{\infty} \mu_\mathcal{T}(\overline{\mathcal{T}}(r_{i,j}, t))dt, \tag{3.5}$$

and for all (i',j') such that $p(i') = (i,j)$,

$$\delta_{(i,j),i'} = \delta_{(i,j),(i',j')} = \int_{t_i}^{t_{i'}} \mu_\mathcal{T}(\overline{\mathcal{T}}(r_{i,j}, t))dt. \tag{3.6}$$

For $K \in \mathbb{N}$, we now build the "first K branches" of the cut-tree. Let $\phi : \mathbb{N} \to \mathcal{E}_\alpha$ denote the bijection such that for all $k \in \mathbb{N}$, $\overline{\mathcal{T}}_{\phi(k)}$ is the k-th biggest tree among the $\overline{\mathcal{T}}_{i,j}$, $(i,j) \in \mathcal{E}_\alpha$, for the order induced by the mass $\mu_\mathcal{T}$. Note that it is well defined, since we consider only the subtrees *above* each cut-point. This bijection ϕ determines the order in which we construct the branches of the cut-tree; note that $\phi(1) = (0,0)$. Let

$$\mathcal{E}(K) = \phi(\{1, \ldots, K\})$$

be the set of the indices corresponding to the first K branches, and

$$\mathcal{I}^*(K) = \{i \in \mathbb{N} : p(i) \in \mathcal{E}(K)\}$$
$$\mathcal{E}^*(K) = (\mathcal{I}^*(K) \times \mathcal{J}_\alpha^+) \setminus \mathcal{E}(K).$$

For all $(i,j) \in \mathcal{E}(K)$, we draw a branch $B_{i,j} = [\![u_{i,j}, v_{i,j}]\!]$ of length $\delta_{(i,j)}$. For all $i \in \mathcal{I}^*(K)$, let b_i be the point of $B_{p(i)}$ which is at distance $\delta_{p(i),i}$ of $u_{p(i)}$. We attach the branches by identifying the points b_i and $u_{i,j}$, for all $(i,j) \in \mathcal{E}(K) \setminus \{(0,0)\}$. This gives an $\mathbb{R}$-tree $\mathrm{Cut}_K(\mathcal{T})$, rooted at $\rho_\mathcal{C} = u_{0,0}$.

We shall prove in the next section that this construction can be done in a consistent way for all $K \in \mathbb{N}$, for example by constructing these $\mathbb{R}$-trees as subspaces of l^1 as in [7]. Thus, we obtain nested trees: $\mathrm{Cut}_K(\mathcal{T}) \subset \mathrm{Cut}_{K'}(\mathcal{T})$ if $K \leq K'$. The cut-tree $\mathrm{Cut}(\mathcal{T})$ is then obtained by taking the limit of $\mathrm{Cut}_K(\mathcal{T})$ as $K \to \infty$.

3.2 Details of the construction

3.2.1 Building the cut-tree as a subset of $l^1_+(\mathbb{N})$

We first introduce some notation for the elements of $l^1_+(\mathbb{N})$. Fix $u = (u(n))_{n\in\mathbb{N}}, v = (v(n))_{n\in\mathbb{N}} \in l^1_+(\mathbb{N})$ (in this subsection only, u denotes en element of $l^1_+(\mathbb{N})$ instead of U), and let $n_{uv} = \max\{n \in \mathbb{N} : u(n) = v(n)\}$. We say that $u < v$ if $u(n_{uv}+1) < v(n_{uv}+1)$ and, for all $n > n_{uv}+1$, $u(n) = 0$. Let $\mathbf{0}$ denote the sequence $(0,0,\ldots)$. Note that any nondecreasing sequence (u_i) such that $d(0,u_i)$ is bounded converges in $l^1_+(\mathbb{N})$. Let

$$u \wedge v = (u(1), \ldots, u(n_{uv}), u(n_{uv}+1) \wedge v(n_{uv}+1), 0, 0, \ldots)$$

be the largest sequence such that $u \wedge v \leq u$ and $u \wedge v \leq v$. Let d be the distance on $l^1_+(\mathbb{N})$ defined by

$$d(\mathbf{0}, u) = \sum_{n\geq 1} u(n) \qquad \text{and} \qquad d(u,v) = d(\mathbf{0},u) + d(\mathbf{0},v) - 2d(\mathbf{0}, u \wedge v).$$

Thus, the shortest path between u and v is

$$[\![u,v]\!] = \left\{w \in l^1_+(\mathbb{N}) : u \wedge v \le w \le u \text{ or } u \wedge v \le w \le v\right\}.$$

Finally, let $\mathbf{1}_N = (\mathbf{1}_{n=N})_{n\in\mathbb{N}}$, and $u_{|N} = (u(1), \ldots, u(N), 0, 0, \ldots)$, for all $N \ge 1$.

We now build the trees $\mathrm{Cut}_K(\mathcal{T})$ as subsets of $l^1_+(\mathbb{N})$, in the following way. We define the points b_i, $u_{i,j}$ and $v_{i,j}$ recursively, by saying that $b_0 = (0, 0, \ldots)$, and for all $(i, j) \in \mathcal{E}_\alpha$, $i' \in \mathbb{N}$ such that $p(i') = (i, j)$,

$$u_{i,j} = b_i, \qquad v_{i,j} = b_i + \delta_{i,j} \cdot \mathbf{1}_{\phi^{-1}(i,j)} \quad \text{and} \quad b_{i'} = b_i + \delta_{(i,j),i'} \cdot \mathbf{1}_{\phi^{-1}(i,j)}$$

Note that the fact that $p(i') = (i, j)$ implies that for any $j' \in \mathcal{J}^+_\alpha$, we have $\phi^{-1}(i, j) < \phi^{-1}(i', j')$: in other words, the branch $B_{i,j}$ is built before $B_{i',j'}$. Finally, let

$$\mathrm{Cut}_K(\mathcal{T}) = \bigcup_{(i,j)\in\mathcal{E}(K)} [\![b_{i,j}, v_{i,j}]\!],$$

and

$$\mathrm{Cut}(\mathcal{T}) = \overline{\bigcup_{K\in\mathbb{N}} \mathrm{Cut}_K(\mathcal{T})},$$

where $\overline{S}$ denotes the closure in (l^1, d) of a subset S.

Before giving the details of our construction, let us make a few remarks on the properties of $\mathrm{Cut}(\mathcal{T})$.

Remark 3.1 (Tree structure of $\mathrm{Cut}(\mathcal{T})$)**.** One can show directly that $\mathrm{Cut}(\mathcal{T})$ is an $\mathbb{R}$-tree. To this end, consider the following condition on a subset S of $l^1_+(\mathbb{N})$:

For all $w, w' \in S$, we have $[\![w, w']\!] \subset S$, and if n is a positive integer such that $w(n), w'(n) \neq 0$, then $n_{w,w'} \ge n - 1$.

This is a sufficient condition for (S, d) to be an $\mathbb{R}$-tree. Moreover, it is elementary to see that these properties are verified for the $\mathrm{Cut}_K(\mathcal{T})$, $K \ge 1$, and that they also hold for the "limit" $\mathrm{Cut}(\mathcal{T})$.

Moreover, the construction allows us to identify the sets of the branching points and leaves of $\mathrm{Cut}(\mathcal{T})$. Indeed, it is clear that

$$\mathcal{B}(\mathrm{Cut}(\mathcal{T})) = \{b_i : i \in \mathbb{N}\} = \{u_e : e \in \mathcal{E}_\alpha \setminus \{(0,0)\}\}, \tag{3.7}$$

and that $b_0 = \mathbf{0}$ is a leaf of $\mathrm{Cut}(\mathcal{T})$. The non-zero leaves of $\mathrm{Cut}(\mathcal{T})$ are the points $w \in \mathrm{Cut}(\mathcal{T})$ such that $\{w' \in \mathrm{Cut}(\mathcal{T}) : w' > w\}$ is empty. Note that for all $N \in \mathbb{N}$, if we let $l^1_N = \left\{w \in l^1_+(\mathbb{N}) : \forall n > N,\ w(n) = 0\right\}$, then for all $K \ge N$, we have $\mathrm{Cut}_K(\mathcal{T}) \cap l^1_N = \mathrm{Cut}_N(\mathcal{T})$, hence

$$\mathrm{Cut}(\mathcal{T}) \cap l^1_N = \mathrm{Cut}_N(\mathcal{T}).$$

This shows that the leaves of $\mathrm{Cut}(\mathcal{T})$ which belong to one of the l^1_N are the v_e, $e \in \mathcal{E}_\alpha$. For all $w \in \mathrm{Cut}(\mathcal{T}) \setminus \bigcup_{N\in\mathbb{N}} l^1_N$, we have $\{w' \in l^1_+(\mathbb{N}) : w' > w\} = \emptyset$, so w is also a leaf of $\mathrm{Cut}(\mathcal{T})$. Thus

$$\mathcal{L}(\mathrm{Cut}(\mathcal{T})) = \left(\mathrm{Cut}(\mathcal{T}) \setminus \bigcup_{N\in\mathbb{N}} l^1_N\right) \cup \{\mathbf{0}\} \cup \{v_e : e \in \mathcal{E}_\alpha\}. \tag{3.8}$$

Remark 3.2 (A sequence of α-stable trees containing the $\mathrm{Cut}_K(\mathcal{T})$)**.** Consider the $\mathbb{R}$-trees $\mathrm{Cut}^*_K(\mathcal{T})$ obtained by grafting, for each $(i, j) \in \mathcal{E}^*(K)$, the root of the tree $\overline{\mathcal{T}}_{i,j}$ at the point b_i of $\mathrm{Cut}_K(\mathcal{T})$. (Note that these trees can be seen as subsets of $l^1_+(\mathbb{N})$ verifying the above condition.) The transformation $\mathcal{T} \mapsto \mathrm{Cut}^*_1(\mathcal{T})$ was studied by Addario-Berry, Broutin and Holmgren [4] in the Brownian case, and by Abraham and Delmas [2] in the general stable case. In particular, they showed that the resulting tree $\mathrm{Cut}_1^*(\mathcal{T})$, endowed with the image measure μ_1 of $\mu_{\mathcal{T}}$, is a standard stable tree, and that $v_{0,0}$, which will be seen later as the leaf of corresponding to the root of $\mathcal{T}$, is a uniform leaf of this tree (in the sense that the *equivalence class* of $((\mathrm{Cut}_1^*(\mathcal{T}), d, \mu_1), \mathbf{0}, v_{0,0})$ has the same distribution as a rooted stable tree with a uniform marked leaf). This result can be used to prove recursively that all the trees $\mathrm{Cut}_K^*(\mathcal{T})$ have the same law as $\mathcal{T}$.

Remark 3.3 ($\mathrm{Cut}(\mathcal{T})$ has finite height)**.** For all $S \subset l_+^1(\mathbb{N})$, let $h(S) = \sup\{d(\mathbf{0}, u) : u \in S\}$; if S is a tree rooted at $\mathbf{0}$, this quantity corresponds to its height. Let us show that $h(\mathrm{Cut}(\mathcal{T}))$ is a.s. finite. It is easy to see that $h(\mathrm{Cut}_K(\mathcal{T})) \to h(\mathrm{Cut}(\mathcal{T}))$ almost surely. Therefore, for all $\eta > 0$, there exists K such that

$$\mathbb{P}\left(|h(\mathrm{Cut}(\mathcal{T})) - h(\mathrm{Cut}_K(\mathcal{T}))| > \eta\right) \leq \eta.$$

For all $h_0 \in \mathbb{R}_+$, we have

$$\begin{aligned}\mathbb{P}\left(h(\mathrm{Cut}(\mathcal{T})) \geq h_0\right) &\leq \mathbb{P}\left(h(\mathrm{Cut}_K(\mathcal{T})) \geq h_0 - \eta\right) + \eta \\ &\leq \mathbb{P}\left(h(\mathrm{Cut}_K^*(\mathcal{T})) \geq h_0 - \eta\right) + \eta \\ &\leq \mathbb{P}\left(h(\mathcal{T}) \geq h_0 - \eta\right) + \eta,\end{aligned}$$

since $\mathrm{Cut}_K^*(\mathcal{T})$ and $\mathcal{T}$ have the same law. Taking $\eta \to 0$, we obtain

$$\mathbb{P}\left(h(\mathrm{Cut}(\mathcal{T})) \geq h_0\right) \leq \mathbb{P}\left(h(\mathcal{T}) \geq h_0\right),$$

hence the conclusion.

In particular, this property has the following useful consequence: any nondecreasing sequence of elements of $\mathrm{Cut}(\mathcal{T})$ converges in $\mathrm{Cut}(\mathcal{T})$.

3.2.2 The correspondence between the points of $\mathcal{T}$ and the leaves of $\mathrm{Cut}(\mathcal{T})$

Lemma 3.4. *There exists a function*

$$\begin{array}{rccc}\ell : & \overline{\mathcal{T}}_{\mathcal{X}} & \to & l_+^1(\mathbb{N}) \\ & x & \mapsto & \ell_x\end{array}$$

such that $\left\{\ell_x : x \in \overline{\mathcal{T}}_{\mathcal{X}}\right\} = \mathcal{L}(\mathrm{Cut}(\mathcal{T})) \cup \mathcal{B}(\mathrm{Cut}(\mathcal{T}))$ *and, for all* $x, y \in \overline{\mathcal{T}}_{\mathcal{X}}$,

$$\begin{aligned}d(\mathbf{0}, \ell_x) &= \int_0^\infty \mu_{\mathcal{T}}(\overline{\mathcal{T}}(x,t))dt, \\ d(\ell_x, \ell_y) &= \int_{t(x,y)}^\infty (\mu_{\mathcal{T}}(\overline{\mathcal{T}}(x,t)) + \mu_{\mathcal{T}}(\overline{\mathcal{T}}(y,t)))dt.\end{aligned}$$

Proof. Let $x \in \overline{\mathcal{T}}_{\mathcal{X}}$. For all $i \geq 1$, let $e_i(x) = \phi(k_i(x))$, with $(k_1(x), \ldots)$ the increasing sequence (finite or infinite) of the k such that $x \in \overline{\mathcal{T}}_{\phi(k)}$. Note that $e_1(x) = (0,0)$, and that the sequences $(\overline{\mathcal{T}}_{e_i(x)})_{i\geq 1}$ and $(t_{e_i(x)})_{i\geq 1}$ are respectively decreasing and increasing. We distinguish three cases:

1. The above sequences are finite, of length $i(x)$, and we have $x = r_{e_{i(x)}}$. This is the case if and only if $x \in \{r_e : e \in \mathcal{E}_\alpha\} = \mathcal{X} \cup \{\rho_{\mathcal{T}}\}$.
2. The above sequences are finite, of length $i(x)$, and we have $x \neq r_{e_{i(x)}}$. This is the case if and only if there exists $e \in \mathcal{E}_\alpha$ such that $x \in \overline{\mathcal{T}}_e$ and $\overline{\mathcal{T}}(x, t_e^+) = \{x\}$. (The fact that such points exist is justified in the remark below.)
3. The above sequences are infinite.

Let S_1, S_2 and S_3 denote the subsets of $\mathrm{Cut}(\mathcal{T})$ corresponding to these three cases (respectively).

Let us briefly comment on the existence of points belonging to S_2. Consider a decreasing sequence (ε_n) in $(0,\infty)$ which converges to zero as $n \to \infty$. Recall that Λ_α denotes the intensity of the Poisson process of cut-points on $\mathbb{R}_+ \times \mathcal{T}$. Since $\Lambda_\alpha((0,\varepsilon_1) \times \mathcal{T}) = \infty$, there exists a cut-point p_1 such that the corresponding cut happens before time ε_1. Recursively, since $\Lambda_\alpha((0,\varepsilon_n) \times \{x \in \mathcal{T} : x \succ p\}) = \infty$ for all $p \in \mathcal{B}(\mathcal{T})$, we can build a sequence of cut-points $(p_n)_{n\geq 1}$ such that for all n, we have $p_{n+1} \succ p_n$, and the cut at point p_n happens before time ε_n. This sequence converges to a point x which will have the property $\overline{\mathcal{T}}(x, 0^+) = \{x\}$. The same argument shows that for all $e \in \mathcal{E}_\alpha$, there exists $x \in \overline{\mathcal{T}}_e$ such that $\overline{\mathcal{T}}(x, t_e^+) = \{x\}$ (iterating the construction "in all branches" from the cut-points p_n would even show that there are uncountably many such points). However, since the total mass of the fragments is a.s. 1 at all times, we have $\mu(S_2) = 0$.

We define the application ℓ by setting

$$\ell_x = \begin{cases} v_{e_{i(x)}(x)} & \text{if } x \in S_1 \\ u_{e_{i(x)}(x)} & \text{if } x \in S_2 \\ \lim_{i\to\infty} u_{e_i(x)} & \text{if } x \in S_3. \end{cases}$$

Note that in the third case, the sequence $(b_{e_i(x)})_{i\geq 1}$ is increasing in $\mathrm{Cut}(\mathcal{T})$, so Remark 3.3 ensures that $\ell_x \in \mathrm{Cut}(\mathcal{T})$. With this definition, we can see that the images of the points in $S_1 \cup S_3$ are the leaves of $\mathrm{Cut}(\mathcal{T})$ (except $\rho_\mathcal{C}$), and the images of the points in S_2 are the branching points of $\mathrm{Cut}(\mathcal{T})$ and its root. Indeed, we have

$$\ell(S_1) = \{\ell_{r_e} : e \in \mathcal{E}_\alpha\} = \{v_e : e \in \mathcal{E}_\alpha\}$$

and

$$\ell(S_2) = \left\{\ell_x : e \in \mathcal{E}_\alpha, x \in \overline{\mathcal{T}}_e \text{ s.t. } \overline{\mathcal{T}}(x, t_e^+) = \{x\}\right\} = \{u_e : e \in \mathcal{E}_\alpha\} = \mathcal{B}(\mathrm{Cut}(\mathcal{T})) \cup \{\mathbf{0}\}$$

(see equation (3.7)). For all $x \in S_3$, we have $\ell_x \in \mathrm{Cut}(\mathcal{T}) \setminus \bigcup_{N\in\mathbb{N}} l_N^1$. Conversely, let $w \in \mathrm{Cut}(\mathcal{T}) \setminus \bigcup_{N\in\mathbb{N}} l_N^1$. For all $N \geq 1$, it is easy to see that $w_{|N} \in \mathrm{Cut}_N(\mathcal{T})$ and that $w_{|N}$ is a branching point of $\mathrm{Cut}(\mathcal{T})$; thus there exists a unique $e[N] \in \mathcal{E}_\alpha$ such that $w_{|N} = b_{e[N]}$. The sequence $(e[N])_{N\geq 1}$ is increasing, so it admits a limit $\in \mathrm{Cut}(\mathcal{T})$. Now the $e_i(x), i \geq 1$ are exactly the $e[N]$, $N \geq 1$, and we have

$$w = \lim_{N\to\infty} w_{|N} = \lim_{N\to\infty} b_{e[N]} = \lim_{i\to\infty} b_{e_i(x)} = \ell_x.$$

Thus $\ell(S_3) = \mathrm{Cut}(\mathcal{T}) \setminus \bigcup_{N\in\mathbb{N}} l_N^1$. Equation(3.8) now shows that

$$\ell(S_1 \cup S_3) = \mathcal{L}(\mathrm{Cut}(\mathcal{T})) \setminus \{\mathbf{0}\},$$

hence

$$\ell(\mathcal{T}) = \mathcal{B}(\mathrm{Cut}(\mathcal{T})) \cup \mathcal{L}(\mathrm{Cut}(\mathcal{T})).$$

We now compute the distances between the points ℓ_x and $\mathbf{0}$ in the three above cases, using the relations (3.5) and (3.6). Let $x \in \mathcal{T}$. Note that for all suitable i, for all $t \in [t_{e_i(x)}, t_{e_{i+1}(x)}]$, $r_{e_i(x)}$ and x are in the same component of the fragmented tree at time t.

— If $x \in S_1$, then we have

$$\begin{aligned} d(\mathbf{0}, \ell_x) &= \sum_{i=1}^{i(x)-1} \delta_{e_i(x), e_{i+1}(x)} + \delta_{e_{i(x)}(x)} \\ &= \sum_{i=1}^{i(x)-1} \int_{t_{e_i(x)}}^{t_{e_{i+1}(x)}} \mu_\mathcal{T}(\overline{\mathcal{T}}(r_{e_i(x)}, t))dt + \int_{t_{e_{i(x)}(x)}}^{\infty} \mu_\mathcal{T}(\overline{\mathcal{T}}(r_{e_{i(x)}(x)}, t))dt \\ &= \sum_{i=1}^{i(x)-1} \int_{t_{e_i(x)}}^{t_{e_{i+1}(x)}} \mu_\mathcal{T}(\overline{\mathcal{T}}(x, t))dt + \int_{t_{e_{i(x)}(x)}}^{\infty} \mu_\mathcal{T}(\overline{\mathcal{T}}(x, t))dt \\ &= \int_0^\infty \mu_\mathcal{T}(\overline{\mathcal{T}}(x, t))dt. \end{aligned}$$

— If $x \in S_2$, then

$$d(\mathbf{0}, \ell_x) = \sum_{i=1}^{i(x)-1} \delta_{e_i(x),e_{i+1}(x)} = \sum_{i=1}^{i(x)-1} \int_{t_{e_i(x)}}^{t_{e_{i+1}(x)}} \mu_{\mathcal{T}}(\overline{\mathcal{T}}(r_{e_i(x)}, t))dt$$
$$= \sum_{i=1}^{i(x)-1} \int_{t_{e_i(x)}}^{t_{e_{i+1}(x)}} \mu_{\mathcal{T}}(\overline{\mathcal{T}}(x,t))dt = \int_0^{\infty} \mu_{\mathcal{T}}(\overline{\mathcal{T}}(x,t))dt,$$

since $\mu_{\mathcal{T}}(\overline{\mathcal{T}}(x,t)) = 0$ for all $t > t_{e_i(x)(x)}$.

— If $x \in S_3$, then

$$d(\mathbf{0}, \ell_x) = \sum_{i \geq 1} \delta_{e_i(x),e_{i+1}(x)} = \sum_{i \geq 1} \int_{t_{e_i(x)}}^{t_{e_{i+1}(x)}} \mu_{\mathcal{T}}(\overline{\mathcal{T}}(x,t))dt,$$

and $t_{e_i(x)} \to \infty$ as $i \to \infty$, hence

$$d(\mathbf{0}, \ell_x) = \int_0^{\infty} \mu_{\mathcal{T}}(\overline{\mathcal{T}}(x,t))dt.$$

Let us finally check the value of $d(\ell_x, \ell_y)$, for $x \neq y \in \overline{\mathcal{T}}_{\mathcal{X}}$. Let $I = \max\{i \geq 1 : e_i(x) = e_i(y)\}$ (note that since $x \neq y$, we have $I < \infty$): for all $i \leq I$, x and y are in the same component of the fragmented tree at time $t_{e_i(x)} = t_{e_i(y)}$. First assume that $I < i(x), i(y)$ (where by convention, $i(x)$ is $+\infty$ if the sequence $(e_i(x))$ is infinite), and let $t = t_{e_{I+1}(x)} \wedge t_{e_{I+1}(y)}$. At time t^-, both x and y are in the same component as $r = r_{e_I(x)}$. If for example $t = t_{e_{I+1}(x)}$, it means that the mark which appears at time t between x and r, separates the components containing x and y (otherwise we would have $e_{I+1}(x) = e_{I+1}(y)$): thus we have $t = t(x,y)$. Besides, we have

$$d(\mathbf{0}, \ell_x \wedge \ell_y) = \sum_{i=1}^{I-1} \delta_{e_i(x),e_{i+1}(x)} + (\delta_{e_I(x),e_{I+1}(x)} \wedge \delta_{e_I(y),e_{I+1}(y)}),$$

and

$$\delta_{e_I(x),e_{I+1}(x)} \wedge \delta_{e_I(y),e_{I+1}(y)} = \left(\int_{t_{e_I(x)}}^{t_{e_{I+1}(x)}} \mu_{\mathcal{T}}(\overline{\mathcal{T}}(r,t))dt\right) \wedge \left(\int_{t_{e_I(y)}}^{t_{e_{I+1}(y)}} \mu_{\mathcal{T}}(\overline{\mathcal{T}}(r,t))dt\right)$$
$$= \int_{t_{e_I(x)}}^{t(x,y)} \mu_{\mathcal{T}}(\overline{\mathcal{T}}(r,t))dt.$$

This yields

$$d(\mathbf{0}, \ell_x \wedge \ell_y) = \sum_{i=1}^{I-1} \int_{t_{e_i(x)}}^{t_{e_{i+1}(x)}} \mu_{\mathcal{T}}(\overline{\mathcal{T}}_{e_i(x)})dt + \int_{t_{e_I(x)}}^{t(x,y)} \mu_{\mathcal{T}}(\overline{\mathcal{T}}(r,t))dt$$
$$= \int_0^{t(x,y)} \mu_{\mathcal{T}}(\overline{\mathcal{T}}(x,t))dt = \int_0^{t(x,y)} \mu_{\mathcal{T}}(\overline{\mathcal{T}}(y,t))dt,$$

hence the expression of $d(\ell_x, \ell_y)$. The cases where $I = i(x)$ or $I = i(y)$ can be treated similarly (for example, if $I = i(x) \neq i(y)$ and $x = r_{e_I(x)}$, we only have to replace $\delta_{e_I(x),e_{I+1}(x)}$ by $\delta_{e_I(x)}$ and t by $t_{e_{I+1}(y)}$ in what precedes; if $I = i(x) \neq i(y)$ and $x \neq r_{e_I(x)}$, we replace $\delta_{e_I(x),e_{I+1}(x)}$ by 0 and t by $t_{e_I(x)}$; etc.). □

To conclude this section, note that the restriction of the correspondence ℓ to $\mathcal{T} \setminus \mathcal{X}$ is measurable. To check this property, it is enough to show that for all $b \in \mathcal{B}(\mathcal{C})$ and $l \in \mathcal{L}(\mathcal{C})$ such that $l \succ b$, the set $\ell^{-1}(\mathcal{C}_b^l) \cap \mathcal{T}$ is a measurable subset of $\mathcal{T}$ (with $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$). The definition of $\mathcal{C}$ ensures that for all such points b and l, there exists $(i,j) \in \mathcal{E}_\alpha$ such that $\mathcal{C}_b^l = \mathcal{C}_{b_i}^{v_{i,j}} = \mathcal{C}_{b_i}^{\ell_{\chi_{i,j}}}$, hence

$$\ell^{-1}(\mathcal{C}_b^l) \cap \mathcal{T} = \overline{\mathcal{T}}(\chi_{i,j}, t_i) \setminus \mathcal{X}.$$

.

3.2.3 Consistency with the previous definitions of the cut-tree

In this section, we justify that the cut-tree we have built is the same as in [19, 28].

Let $(L_n, n \in \mathbb{N})$ be independent uniform leaves of $\mathcal{T}$, $l_n = \ell_{L_n}$ for all $n \in \mathbb{N}$, and $l_0 = \rho_{\mathcal{C}}(= \mathbf{0})$. The above lemma shows that for all $n \geq 1$, we have

$$d(l_0, l_n) = \int_0^\infty \mu_{\mathcal{T}}(\overline{\mathcal{T}}(L_n, t))dt,$$

and for all $n, n' \geq 1$, we have

$$d(l_n, l_{n'}) = \int_{t(L_n, L_{n'})}^\infty (\mu_{\mathcal{T}}(\overline{\mathcal{T}}(L_n, t)) + \mu_{\mathcal{T}}(\overline{\mathcal{T}}(L_{n'}, t)))dt.$$

Thus, for example in the case $\alpha = 2$, for all $k \in \mathbb{N}$, the matrix of the distances $(d(l_n, l_{n'}))_{0 \leq n, n' \leq k}$ is exactly the same as in the definition of the cut-tree given by Bertoin and Miermont in [19]. As a consequence, these two definitions of the cut-tree coincide, and yield a standard Brownian tree (see [19, Theorem 1]). Similarly, in the case $\alpha \in (1, 2)$, the cut-tree we constructed here is the same as in [28], and Theorem 1.3 of [28] shows that it is a standard stable tree of index α.

As a consequence, to conclude the proof of Proposition 1.3, we only have to prove the following lemma:

Lemma 3.5. *Almost surely, we have* $\{l_n, n \geq 1\} \subset \mathcal{L}(\mathrm{Cut}(\mathcal{T}))$*, and*

$$\mathrm{Cut}(\mathcal{T}) = \overline{\bigcup_{n \in \mathbb{N}} [\![\mathbf{0}, l_n]\!]}.$$

Proof. All the properties stated here hold almost surely. First note the following facts:

- for all $e \in \mathcal{E}_\alpha$, since $\mu_{\mathcal{T}}(\overline{\mathcal{T}}_e) > 0$, there exists $n \geq 0$ such that $L_n \in \overline{\mathcal{T}}_e$ (and by construction, $u_e < l_n$);
- for all $n, n' \in \mathbb{N}$, $d(\mathbf{0}, l_{n'}) + d(l_{n'}, l_n) > d(\mathbf{0}, l_n)$ (cf. the above remark on the distribution of $(d(l_n, l_{n'}))_{n, n' \geq 1}$), so we do not have $l_{n'} < l_n$.

As a consequence, we have $\{L_n : n \in \mathbb{N}\} \cap S_2 = \emptyset$. (In fact, it is even a subset of S_3, since $S_1 \subset \mathcal{X}$ and $\mu_{\mathcal{T}}(\mathcal{X}) = 0$.) Thus the points l_n, $n \in \mathbb{N}$ are leaves of $\mathrm{Cut}(\mathcal{T})$.

To complete the proof, it is enough to show that $\mathcal{B}(\mathrm{Cut}(\mathcal{T}))$ is dense in $\mathrm{Cut}(\mathcal{T})$, and that it is contained in $\bigcup_{n \in \mathbb{N}} [\![\mathbf{0}, l_n]\!]$. For all $i \in \mathbb{N}$, there exists n, n' such that $\chi_i \in [\![L_n, L_{n'}]\!]$, which means that $b_i = l_n \wedge l_{n'}$, hence the second point. Besides, it has already been noted in the previous proof that any point $w \in \mathrm{Cut}(\mathcal{T}) \setminus \bigcup_{N \in \mathbb{N}} l_N^1$ can be approached by the branching points $w_{|N}$, $N \in \mathbb{N}$. Now, for all $N \in \mathbb{N}$, we have $\mathrm{Cut}(\mathcal{T}) \cap l_N^1 = \mathrm{Cut}_N(\mathcal{T})$. Recall Remark 3.2. By construction, the branching points of $\mathrm{Cut}(\mathcal{T})$ and $\mathrm{Cut}_N^*(\mathcal{T})$ in l_N^1 are the same:

$$\mathcal{B}(\mathrm{Cut}(\mathcal{T})) \cap l_N^1 = \mathcal{B}(\mathrm{Cut}_N^*(\mathcal{T})) \cap l_N^1.$$

Since $\mathrm{Cut}_N^*(\mathcal{T})$ is a stable tree, $\mathcal{B}(\mathrm{Cut}_N^*(\mathcal{T}))$ is dense in $\mathrm{Cut}_N^*(\mathcal{T})$, and *a fortiori* in $\mathrm{Cut}_N(\mathcal{T})$. Noting that for all $b \in \mathcal{B}(\mathrm{Cut}_N^*(\mathcal{T}))$ and $w \in \mathrm{Cut}_N(\mathcal{T})$, we have

$$d(b_{|N}, w) \leq d(b, w)$$

and $b_{|N} \in \mathcal{B}(\mathrm{Cut}_N^*(\mathcal{T})) \cap l_N^1 \subset \mathcal{B}(\mathrm{Cut}(\mathcal{T}))$, we get the density of $\mathcal{B}(\mathrm{Cut}(\mathcal{T}))$ in $\mathrm{Cut}_N(\mathcal{T})$. □

Note that this Lemma also shows that the equivalence class of $\mathrm{Cut}(\mathcal{T})$ does not depend on the root of $\mathcal{T}$, even if we first used $\rho_{\mathcal{T}}$ as a starting point for our construction. This explains why we need a distinguished leaf of $\mathrm{Cut}(\mathcal{T})$ to keep track of the position of $\rho_{\mathcal{T}}$.

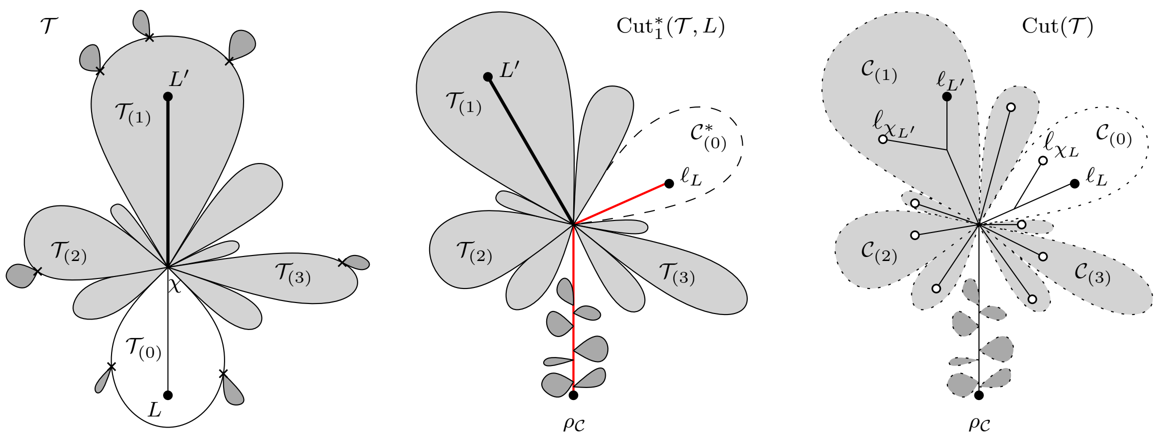

Figure 3.2 – The leftmost picture shows the components $(\mathcal{T}_{(j)}, j \in \mathcal{J}_\alpha)$ created from the subtree $\overline{\mathcal{T}}(L, t(L, L')^-)$ when we cut at point $\chi = \chi(L, L')$; the fragments in dark grey are the ones which are separated from the root before time $t(L, L')$, so they do not appear in the trees $\mathcal{T}_{(j)}$. These components can naturally be seen as subtrees of $\mathrm{Cut}_1^*(\mathcal{T}, L)$: the ones in dark grey are grafted along the first branch before the branching point B, and the $\mathcal{T}_{(j)}$, $j \geq 1$ (in light grey), correspond to the subtrees above B (except the one containing ℓ_L). The last picture illustrates the fact that the cut-trees of each of these components appear as subtrees of $\mathrm{Cut}(\mathcal{T})$. The points indicated by white vertices are the points of $f(b)$, with $b = \ell_L \wedge \ell_{L'}$.

4 The case where $\mathcal{C}$ is the cut-tree of a stable tree

In this whole section, we work under the hypothesis that $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$ for a stable tree $\mathcal{T}$ of index α. Our main goal is to prove the identity stated in Proposition 1.7:

$$\delta_{\mathrm{Cut}(\mathcal{T})}(\ell_L, \ell_{L'}) = d_{\mathcal{T}}(L, L') \qquad \text{a.s.},$$

for independent uniform leaves L and L' of $\mathcal{T}$. We will also show that the decoration $f^{\mathcal{T}}$ verifies Condition 1.4. We begin with a technical lemma which will be used to show both results.

4.1 Joint distribution of the components created by the first cut between two uniform leaves of $\mathcal{T}$

Let L and L' denote two independent uniform leaves of $\mathcal{T}$. Recall that $t(L, L')$ denotes the first time at which a cut-point appears on the segment $[\![L, L']\!]$. We are interested in the joint distribution of the trees into which $\overline{\mathcal{T}}(L, t(L, L')^-)$ is split at time $t(L, L')$. To simplify notation, we relabel them in the following way, as shown on the first picture of Figure 3.2:

— $\mathcal{T}_{(0)}$ is the component containing L (equal to $\overline{\mathcal{T}}(L, t(L, l')))$,

— $\mathcal{T}_{(1)}$ is the component containing L' (equal to $\overline{\mathcal{T}}(L', t(L, l')))$,

— if $\alpha \neq 2$, the remaining components $\mathcal{T}_{(j)}$, for $j \geq 2$, are ordered by decreasing order of mass. (For $\alpha = 2$ there are only two components.)

We see the $\mathcal{T}_{(j)}$, $j \in \mathcal{J}_\alpha$ as random rooted trees (rooted at the cut-point $\chi(L, L')$). Moreover, we let $\mathcal{T}^\bullet_{(0)}$ denote the rooted tree with a distinguished leaf $(\mathcal{T}_{(0)}, L)$, and similarly $\mathcal{T}^\bullet_{(1)} = (\mathcal{T}_{(1)}, L')$. We also let χ_L and $\chi_{L'}$ denote the points in $\{\chi_{i,j} : (i,j) \in \mathbb{N} \times \mathcal{J}_\alpha\}$ which correspond to the cut-point $\chi(L, L')$ "seen from" $\overline{\mathcal{T}}(L, t(L, L'))$ and $\overline{\mathcal{T}}(L', t(L, L'))$ respectively.

Lemma 4.1. *Conditionally on their masses, the random trees $(\mathcal{T}^\bullet_{(0)}, \mathcal{T}^\bullet_{(1)}, \mathcal{T}_{(j)}, j \geq 2)$ are mutually independent, with the following distributions:*

— *$\mathcal{T}^\bullet_{(0)}$ and $\mathcal{T}^\bullet_{(1)}$ are α-stable rooted trees with a uniform distinguished leaf.*

— *if $\alpha \neq 2$, for all $j \geq 2$, $\mathcal{T}_{(j)}$ is an α-stable rooted tree.*

Proof. Recall from Section 3 the tree $\mathrm{Cut}_1^*(\mathcal{T})$ which is obtained by building the first branch $[\![\rho_\mathcal{C}, \ell_{\rho_\mathcal{T}}]\!]$ of the cut-tree, together with the points b_i, $i \in \mathcal{I}^*(1)$, and grafting the root of the tree $\overline{\mathcal{T}}_{i,j}$ at point b_i, for all $(i,j) \in \mathcal{E}^*(1)$. We let $\mathrm{Cut}_1^*(\mathcal{T}, L)$ denote the tree obtained by performing this construction after rerooting the tree $\mathcal{T}$ at point L (so that the "first branch" we build is $[\![\rho_\mathcal{C}, \ell_L]\!]$). Since the distribution of $\mathcal{T}$ is invariant under uniform rerooting, we can apply the results of [4] to $\mathrm{Cut}_1^*(\mathcal{T}, L)$.

Theorem 1.5 of [4, 2] shows that the tree $\mathrm{Cut}_1^*(\mathcal{T}, L)$, endowed with the image measure μ_1 of $\mu_\mathcal{T}$, is a standard stable tree, and that ℓ_L is a uniform leaf of this tree. Moreover, L' can be seen as a point of $\mathrm{Cut}_1^*(\mathcal{T}, L)$, and since μ_1 is the image measure of $\mu_\mathcal{T}$, it is a uniform leaf of $\mathrm{Cut}_1^*(\mathcal{T}, L)$, independent of ℓ_L.

Let B denote the most recent common ancestor of ℓ_L and L' in $\mathrm{Cut}_1^*(\mathcal{T}, L)$. Theorem 2.1 gives the joint distribution of the subtrees of $\mathrm{Cut}_1^*(\mathcal{T}, L)$ which are above B. With the natural identifications between these components and the $\mathcal{T}_{(i)}, i \geq 1$ (as shown on Figure 3.2), we get that conditionally on their masses:

— $\mathcal{T}_{(1)}^\bullet = ((\mathrm{Cut}_1^*(\mathcal{T}, L))_B^{L'}, L')$ is a rooted α-stable tree in which L' is a uniform leaf,

— if $\alpha \neq 2$, for all $j \geq 2$, $\mathcal{T}_{(j)}$ is a rooted α-stable tree,

— $\mathcal{C}_{(0)}^* := (\mathrm{Cut}_1^*(\mathcal{T}, L))_B^{\ell_L}$ is a rooted α-stable tree in which ℓ_L is a uniform leaf,

and these random trees are mutually independent.

To conclude the proof, we need to see that we can recover the tree $\mathcal{T}_{(0)}^\bullet$ by a deterministic transformation of $(\mathcal{C}_{(0)}^*, \ell_L, \chi_L)$ (note that χ_L can be seen as a uniform random point in this tree). To this end, we use the inverse transformation of Cut_1^*, also studied in [4]. Indeed, we have $\mathcal{C}_{(0)}^* = \mathrm{Cut}_1^*(\mathcal{T}_{(0)}, L)$ for the fragmentation given by the cut-points falling into $\mathcal{T}_{(0)}$ (after time $t(L, L')$). By applying the inverse transformation to $\mathcal{C}_{(0)}^*$, we thus get a tree $\widetilde{\mathcal{T}}_{(0)}$ which can be seen as $\mathcal{T}_{(0)}$ rerooted at point L, and we can also recover the point χ_L: this gives us all the information we need on $\mathcal{T}_{(0)}^\bullet$. In particular, it shows that $\mathcal{T}_{(0)}^\bullet$ is independent of the trees $\mathcal{T}_{(1)}^\bullet$ and $\mathcal{T}_{(i)}$, $i \geq 2$.

Finally, since $(\mathcal{T}, L, L')$ and $(\mathcal{T}, L', L)$ have the same distribution, we have that

$$(\mathcal{T}_{(0)}^\bullet, \mathcal{T}_{(1)}^\bullet, \mathcal{T}_{(2)}, \mathcal{T}_{(3)}, \ldots) \stackrel{(d)}{=} (\mathcal{T}_{(1)}^\bullet, \mathcal{T}_{(0)}^\bullet, \mathcal{T}_{(2)}, \mathcal{T}_{(3)}, \ldots),$$

hence the fact that χ_L is a uniform leaf of $\mathcal{T}_{(0)}$. $\square$

The first important consequence of Lemma 4.1 is the following:

Corollary 4.2. *The decoration $f^\mathcal{T}$ verifies Condition 1.4.*

Proof. First note that it is enough to check that the properties stated in Condition 1.4 hold for $b = l \wedge l'$, where $l = \ell_L$ and $l' = \ell_{L'}$ for two independent uniform leaves L and L' of $\mathcal{T}$, independent of the fragmentation (so that l and l' are independent of $f^\mathcal{T}$). As shown in Figure 3.2, we write $\mathcal{C}_{(0)} = \mathcal{C}_b^l$, $\mathcal{C}_{(1)} = \mathcal{C}_b^{l'}$, and let $\mathcal{C}_{(j)}$, $j \in \mathcal{J}_\alpha \setminus \{0, 1\}$ denote the remaining connected components of $\mathcal{C} \setminus \{b\}$, except the one containing the root, ordered by decreasing order of mass. We number the elements of $f(b)$ accordingly, by writing $f^{(j)}(b)$ for the unique element of $\mathcal{C}_{(j)} \cap f(b)$. In the whole proof, we work conditionally on the masses of the components $\mathcal{C}_{(j)}$, $j \geq 0$. We have that:

— $\mathcal{C}_{(0)}$ can be seen as the cut-tree of $\mathcal{T}_{(0)}$, with l and $f^{(0)}(b)$ the leaves corresponding (respectively) to the root χ_L and the point L,

— $\mathcal{C}_{(1)}$ can be seen as the cut-tree of $\mathcal{T}_{(1)}$, with l' and $f^{(1)}(b)$ the leaves corresponding (respectively) to the root $\chi_{L'}$ and the leaf L',

— for all $j \in \mathcal{J}_\alpha \setminus \{0, 1\}$, $\mathcal{C}_{(j)}$ can be seen as the cut-tree of $\mathcal{T}_{(j)}$, with $f^{(j)}(b)$ the leaf corresponding to the root.

Putting together Proposition 1.3 and Lemma 4.1, we get the following joint distribution:

— $(\mathcal{C}_{(0)}, l, f^{(0)}(b))$ and $(\mathcal{C}_{(1)}, l', f^{(1)}(b))$ are rooted α-stable trees with two uniform independent leaves,

— for all $j \in \mathcal{J}_\alpha \setminus \{0, 1\}$, $(\mathcal{C}_{(j)}, f^{(j)}(b))$ is a rooted α-stable tree with a uniform leaf,

— these trees (with their distinguished points) are mutually independent.

Besides, the Poisson point processes of marks on each of the subtrees $\mathcal{T}_{(j)}$, $j \in \mathcal{J}_\alpha$ are mutually independent, and independent of χ. As a consequence, the restrictions of $f^{\mathcal{T}}$ to the subtrees $\mathcal{C}_{(j)}$, $j \in \mathcal{J}_\alpha$ are mutually independent, and independent of $f(b)$. □

4.2 Proof of Proposition 1.7

As in the previous section, we let L and L' be two independent uniform leaves of $\mathcal{T}$. Recall the construction of $\mathcal{C}^{\ell_L,\ell_{L'}}$ introduced in Section 1.2.4, and especially the sequences $l(u)$, $l'(u)$ and $\mathcal{C}(u)$, $u \in U$. We now build sequences which will encode the corresponding points and components of $\mathcal{T}$.

For all $x, y \in \mathcal{T}$ such that $\ell_x, \ell_y \in \mathcal{L}(\mathcal{C})$, we use the notation $\chi_x(x,y) = \chi_x(y,x)$ for the representative point of $\chi(x,y)$ in $\overline{\mathcal{T}}(x, t(x,y))$:

$$\chi_x(x,y) = \chi_x(y,x) := \chi_{i,j} \qquad \text{if } \chi(x,y) = \chi_i \text{ and } \chi_{i,j} \in \overline{\mathcal{T}}(x, t(x,y)).$$

The sequences are defined as follows:

— Let $L(\emptyset) = L$, $L'(\emptyset) = L'$, and $\mathcal{T}(\emptyset) = \mathcal{T}$.

— For all $u \in U$,

$$\begin{aligned} &L(u,1) = L(u) \qquad L'(u,1) = \chi_{L(u)}(L(u), L'(u)) \qquad \mathcal{T}(u,1) = \overline{\mathcal{T}}(L(u), t(L(u), L'(u))) \\ &L(u,2) = \chi_{L'(u)}(L(u), L'(u)) \qquad L'(u,2) = L'(u) \qquad \mathcal{T}(u,2) = \overline{\mathcal{T}}(L'(u), t(L(u), L'(u))). \end{aligned}$$

It is clear that for all $u \in U$, we have $l(u) = \ell_{L(u)}$, $l'(u) = \ell_{L'(u)}$, and $\mathcal{C}(u) = \mathrm{Cut}(\mathcal{T}(u))$. In particular, $m(u) = \mu_{\mathcal{C}}(\mathcal{C}(u)) = \mu_{\mathcal{T}}(\mathcal{T}(u))$.

Our goal is to prove that we have $d_{\mathcal{T}}(L, L') = \delta(\ell_L, \ell_{L'})$. A first interesting similarity between these two quantities is the fact that for all $n \geq 1$, we have

$$d_{\mathcal{T}}(L, L') = \sum_{u \in \{1,2\}^n} d_{\mathcal{T}}(L(u), L'(u)) \qquad \text{and} \qquad \delta(\ell_L, \ell_{L'}) = \sum_{u \in \{1,2\}^n} \delta(\ell_{L(u)}, \ell_{L'(u)}).$$

We will not exactly use this property, but rather the fact that $\delta(\ell_L, \ell_{L'})$ is a multiple of the limit of $W_n = \sum_{u \in \{1,2\}^n} m(u)^{1-1/\alpha}$. The key argument to our proof is the following:

Lemma 4.3. *Fix $n \geq 1$. The trees $\mathcal{T}(u), u \in \{1,2\}^n$ are independent stable trees of index α, with respective masses $m(u), u \in \{1,2\}^n$. Conditionally on these trees, for all $u \in \{1,2\}^n$, $L(u)$ and $L'(u)$ are independent uniform leaves of $\mathcal{T}(u)$. In particular, we have the identity*

$$d_{\mathcal{T}}(L, L') = \sum_{u \in \{1,2\}^n} m(u)^{1-1/\alpha} \Delta_u,$$

where $(\Delta_u, u \in \{1,2\}^n)$ are i.i.d. variables having the same distribution as $d_{\mathcal{T}}(L, L')$, independent of $(m(u), u \in \{1,2\}^n)$.

For $n = 1$, this is a direct application of Lemma 4.1; we obtain the result for $n \geq 2$ by iterating.

Let us now prove Proposition 1.7. We write $D = d_{\mathcal{T}}(L, L')$, and recall that $\delta(L, L') = c_\alpha \cdot W_\infty$. Since D/W_n converges a.s. to D/W_∞ as $n \to \infty$, it is enough to show that there exists a subsequence of D/W_n which converges in probability to c_α. Using the notation of Lemma 4.3, we have

$$\frac{D}{W_n} = \sum_{u \in \{1,2\}^n} \lambda_u \Delta_u,$$

where $\lambda_u = m(u)^{1-1/\alpha}/W_n$ for all $u \in U$. Note that

$$\sum_{u\in\{1,2\}^n} \lambda_u = \frac{1}{W_n} \sum_{u\in\{1,2\}^n} m(u)^{1-1/\alpha} = 1.$$

We compute the expectation and the variance of D/W_n:

$$\mathbb{E}\left[\frac{D}{W_n}\right] = \mathbb{E}\left[\sum_{u\in\{1,2\}^n} \lambda_u \mathbb{E}\left[\Delta_u | \lambda_v, v \in \{1,2\}^n\right]\right] = \mathbb{E}\left[\sum_{u\in\{1,2\}^n} \lambda_u \cdot c_\alpha\right] = c_\alpha,$$

and

$$\begin{aligned}
\mathrm{Var}\left(\frac{D}{W_n}\right) &= \mathbb{E}\left[\sum_{u,v\in\{1,2\}^n} \lambda_u\lambda_v\Delta_u\Delta_v\right] - c_\alpha^2 \\
&= \mathbb{E}\left[\sum_{u,v\in\{1,2\}^n} \lambda_u\lambda_v\mathbb{E}\left[\Delta_u\Delta_v | \lambda_w, w \in \{1,2\}^n\right]\right] - c_\alpha^2 \\
&= c_\alpha^2\mathbb{E}\left[\sum_{u,v\in\{1,2\}^n} \lambda_u\lambda_v\right] + \mathrm{Var}(\Delta)\mathbb{E}\left[\sum_{u\in\{1,2\}^n} \lambda_u^2\right] - c_\alpha^2,
\end{aligned}$$

with $\Delta \overset{(d)}{=} \Delta_u$ for all u. Since $\sum_{u,v\in\{1,2\}^n} \lambda_u\lambda_v = \left(\sum_{u\in\{1,2\}^n} \lambda_u\right)^2 = 1$ a.s., this gives

$$\mathrm{Var}\left(\frac{D}{W_n}\right) = \mathrm{Var}(\Delta)\mathbb{E}\left[\sum_{u\in\{1,2\}^n} \lambda_u^2\right].$$

Let $Z_n = \sum_{u\in\{1,2\}^n} m(u)^{2-2/\alpha}$, so that

$$\mathrm{Var}\left(\frac{D}{W_n}\right) = \mathrm{Var}(\Delta)\mathbb{E}\left[\frac{Z_n}{W_n^2}\right].$$

For all $n \in \mathbb{N}$, we have

$$\begin{aligned}
\mathbb{E}\left[Z_{n+1}\right] &= \mathbb{E}\left[\sum_{u\in\{1,2\}^n} m(u)^{2-2/\alpha}\frac{m(u,1)^{2-2/\alpha} + m(u,2)^{2-2/\alpha}}{m(u)^{2-2/\alpha}}\right] \\
&= \mathbb{E}\left[\sum_{u\in\{1,2\}^n} m(u)^{2-2/\alpha}\mathbb{E}\left[m(1)^{2-2/\alpha} + m(2)^{2-2/\alpha}\right]\right] \\
&= \frac{2}{3}\mathbb{E}\left[Z_n\right],
\end{aligned}$$

so $\mathbb{E}\left[Z_n\right] = (2/3)^n$. Thus Z_n converges in L^1 towards 0. Note that this does not imply directly that $\mathbb{E}\left[Z_n/W_n^2\right]$ converges to 0: we have to use an almost sure convergence as an intermediate. Letting ϕ be an extraction such that $Z_{\phi(n)}$ converges almost surely as $n \to \infty$, we get that

$$\sum_{u\in\{1,2\}^{\phi(n)}} \lambda_u^2 = \frac{Z_{\phi(n)}}{W_{\phi(n)}^2} \xrightarrow[n\to\infty]{} 0 \qquad \text{a.s.}$$

Now, since this sequence is dominated by 1, the convergence also holds in L^1. As a consequence, $\mathrm{Var}(D/W_{\phi(n)})$ converges to 0 as $n \to \infty$. The convergence in probability of $D/W_{\phi(n)}$ is now a straightforward consequence of the Chebychev inequality.

5 Correspondence between the points of $\mathcal{C}$ and $\mathrm{Rec}(\mathcal{C})$ and fragmentation of $\mathrm{Rec}(\mathcal{C})$

In this section, we consider a sequence l_i, $i \in \mathbb{N}$ of independent uniform leaves of $\mathcal{C}$, independent of $l_0 := l^0$, and let $\mathcal{D} := \{l_i : i \in \mathbb{N}\}$. Since the root $\rho_{\mathcal{C}}$ does not play the same role as the other leaves when $\mathcal{C}$ is the cut-tree of a stable tree, we also introduce the notation $\mathcal{L}^\bullet(\mathcal{C}) = \mathcal{L}(\mathcal{C}) \setminus \{\rho_{\mathcal{C}}\}$.

5.1 Continuous extension of the reconstructed distance to $\mathcal{L}^\bullet(\mathcal{C})$

We use the fact that the set $\mathcal{D}$ is a.s. dense in $\mathcal{C}$ to define the reconstructed distance $\delta(x, y)$ for any points $x, y \in \mathcal{L}^\bullet(\mathcal{C})$. The main idea consists in showing the following property:

Proposition 5.1. *Almost surely, for all $x \in \mathcal{L}^\bullet(\mathcal{C})$ and $\varepsilon > 0$, there exists a neighbourhood $V_x(\varepsilon)$ of x such that*

$$\sup_{l,l' \in \mathcal{D} \cap V_x(\varepsilon)} \delta(l, l') \leq \varepsilon. \tag{3.9}$$

Let us first explain how this allows us to extend δ to $\mathcal{L}^\bullet(\mathcal{C})$. For any points $x, y \in \mathcal{L}(\mathcal{C})$, let

$$R_{x,y}(\varepsilon) = \left\{\delta(l, l') : l \in \mathcal{D} \cap V_x(\varepsilon), l' \in \mathcal{D} \cap V_y(\varepsilon)\right\},$$

and

$$R_{x,y} = \bigcap_{\varepsilon > 0} \overline{R_{x,y}(\varepsilon)},$$

where $\overline{S}$ denotes the closure of a subset S of $\mathbb{R}$. The set $R_{x,y}$ is non-empty, since it is written as a decreasing intersection of non-empty compact subsets. Besides, equation (3.9) shows that for all ε, the diameter of $R_{x,y}(\varepsilon)$ is less than ε. As a consequence, $R_{x,y}$ has a unique element, which we call the reconstructed distance $\delta(x, y)$. Note that we do not claim that $\delta(x, y)$ can be computed as the limit of the rescaled quantities $N^{x,y}(\varepsilon)$.

The proof of Proposition 5.1 will be divided into two steps. We first show that we can split $\mathcal{C}$ into stable subtrees on which δ is small enough, and that these subtrees can be used to define $V_x(\varepsilon)$ for all but a countable number of points in $\mathcal{L}^\bullet(\mathcal{C})$. This construction bears some similarities with that of Section 3.

For any subtree $\mathcal{S}$ of $\mathcal{C}$, endowed with a distinguished leaf $l^* \in \mathcal{L}(\mathcal{C}) \cap \mathcal{S}$, let

$$\mathcal{F}_0(\mathcal{S}, l^*) = \left\{\left(\mathcal{S}^l_{l \wedge l^*}, f_l(l \wedge l^*)\right) : l \in \mathcal{D} \cap \mathcal{S}\right\}.$$

In particular, $\mathcal{F}_0(\mathcal{C}, l^0)$ is the set of the connected components of $\mathcal{C} \setminus [\![\rho_{\mathcal{C}}, l^0]\!]$, with a distinguished leaf for each component. The joint distribution of their masses, ranked in decreasing order, is known from [38, Corollary 10] (it corresponds to the fine spinal mass partition), and conditionally on these masses, the components are independent stable trees of index α. Recursively, we define $\mathcal{F}_K(\mathcal{C}, l^0)$ by letting $(\mathcal{S}_K, l^K)$ be the element of $\mathcal{F}_K(\mathcal{C}, l^0)$ which has the biggest mass, and setting

$$\mathcal{F}_{K+1}(\mathcal{C}, l^0) = \left(\mathcal{F}_K(\mathcal{C}, l^0) \setminus \left\{(\mathcal{S}_K, l^K)\right\}\right) \cup \mathcal{F}_0(\mathcal{S}_K, l^K).$$

Again, conditionally on their masses, the elements of $\mathcal{F}_K(\mathcal{C}, l^0)$ are independent stable trees of index α, with each a uniform distinguished leaf and a suitable decoration. Also note that $\mu_{\mathcal{C}}(\mathcal{S}_K)$ almost surely converges to 0 as $k \to \infty$; as a consequence, for all $b \in \mathcal{B}(\mathcal{C})$ and $l > b$, there exists K such that $\mu_{\mathcal{C}}(\mathcal{C}^l_b) > \mu_{\mathcal{C}}(\mathcal{S}_K)$, which implies that $b \in \bigcup_{0 \leq k \leq K} [\![\rho_{\mathcal{C}}, l^k]\!]$.

Lemma 5.2. *Almost surely, for all $\varepsilon > 0$, there exists a random integer $K(\varepsilon)$ such that*

$$\sup_{(\mathcal{S}, l^*) \in \mathcal{F}_K(\mathcal{C}, l^0)} \sup_{l,l' \in \mathcal{S} \cap \mathcal{D}} \delta(l, l') \leq \varepsilon.$$

Proof. Consider the two quantities

$$\Delta_K = \sup_{(\mathcal{S},l^*)\in\mathcal{F}_K(\mathcal{C},l^0)} \sup_{l,l'\in\mathcal{S}\cap\mathcal{D}} \delta(l,l')$$
$$D_K = \sup_{(\mathcal{S},l^*)\in\mathcal{F}_K(\mathcal{C},l^0)} \sup_{l,l'\in\mathcal{S}\cap\mathcal{D}} d_{\mathcal{C}}(l,l').$$

Identity (3.2) and the above remarks on the joint distribution of the $(\mathcal{S},l^*)\in\mathcal{F}_K(\mathcal{C},l^0)$ show that for all $K\in\mathbb{N}$, D_K and Δ_K have the same distribution. Thus it is enough to show that D_K converges almost surely to 0 as $K\to\infty$: indeed, this implies that Δ_K converges in probability towards 0, and since $(\Delta_K)_{K\geq 0}$ is decreasing, the convergence also holds almost surely.

To prove the almost sure convergence of D_K, we use the compactness of $\mathcal{C}$, and the fact that $\mathcal{B}(\mathcal{C})$ is dense in $\mathcal{C}$ and contained in $\bigcup_{k\geq 0}[\![\rho_{\mathcal{C}},l^k]\!]$. Let $\varepsilon>0$. Almost surely, there exists a covering of $\mathcal{C}$ with a finite number of balls of radius $\varepsilon/2$, centered at points of $\bigcup_{0\leq k\leq K}[\![\rho_{\mathcal{C}},l^k]\!]$ for a large enough K. This implies $D_k\leq\varepsilon$ for all $k\geq K$. □

Note that for all $K\in\mathbb{N}$, we have

$$\bigcup_{(\mathcal{S},l^*)\in\mathcal{F}_K(\mathcal{C},l^0)} \mathcal{L}^\bullet(\mathcal{S}) = \mathcal{L}^\bullet(\mathcal{C})\setminus\left\{l^0,\ldots,l^{K-1}\right\}.$$

Therefore, for all $x\in\mathcal{L}^\bullet(\mathcal{C})\setminus\left\{l^k : k\geq 0\right\}$, we can define the neighbourhood $V_x(\varepsilon)$ as the unique element of $\mathcal{F}_{K(\varepsilon)}(\mathcal{C},l^0)$ containing x. The above lemma shows that equation (3.9) is verified on $V_x(\varepsilon)$.

The second (and last) step of the proof of Proposition 5.1 consists in justifying the existence of the neighbourhoods $V_{l^k}(\varepsilon)$ for all $k\geq 0$ and $\varepsilon>0$. Let $k\in\mathbb{N}$. Conditionally on the mass $M_k=\mu_{\mathcal{C}}(\mathcal{S}_k)$, the space $(\mathcal{S}_k, d_{\mathcal{C}}/M_k^{1-1/\alpha}, \mu_{\mathcal{C}}/M_k)$ is a stable tree of index α, on which the restriction of the decoration f and the distinguished point l^k allow us to define a reconstructed distance $\delta_{\mathcal{S}_k}$. In particular, almost surely, the distances $\delta_{\mathcal{S}_k}(l,l^k)$, for $l\in\mathcal{D}\cap\mathcal{S}_k$, are well-defined. Moreover, it can be seen from the definition of $\delta_{\mathcal{C}}$ that for all $l,l'\in(\mathcal{D}\cap\mathcal{S}_k)\cup\left\{l^k\right\}$, we have

$$\delta_{\mathcal{C}}(l,l') = M_k^{1-1/\alpha}\delta_{\mathcal{S}_k}(l,l'). \tag{3.10}$$

(Indeed, if we let $N^{l,\tilde{l}}_{\mathcal{S}_k}(\varepsilon)$ denote the quantity obtained by performing the construction of Section 1.2.4 in the tree $\mathcal{S}_k$ rescaled to have mass 1, we have $N^{l,l'}_{\mathcal{S}_k}(\varepsilon) = N^{l,l'}(\varepsilon/M_k)$.) Also note that if $l,\tilde{l}$ are respectively uniform leaves of $\mathcal{S}_k$ and $\mathcal{C}$, then we have

$$\delta_{\mathcal{S}_k}(l,l^k) \overset{(d)}{=} \delta_{\mathcal{C}}(\tilde{l},l^0).$$

The two above equations imply that it is enough to show the existence of the neighbourhoods $V_{l^0}(\varepsilon)$ such that (3.9) holds, for all $\varepsilon>0$. This is a straightforward consequence of the following lemma:

Lemma 5.3. *Let $\varepsilon>0$. Almost surely, there exists a neighbourhood $V_{l^0}(\varepsilon)$ such that*

$$\sup_{\tilde{l}\in\mathcal{D}\cap V_{l^0}(\varepsilon)} \delta(l^0,\tilde{l}) \leq \varepsilon/2.$$

Note that this choice ensures that for all $k\geq 0$, for all $l\in\mathcal{D}$, the two possible definitions of $\delta(l^k,l)$ (using equation (3.10) or the construction explained at the beginning of the section) coincide.

Proof. We perform the construction of Section 1.2.4 for $l=l^0$ and an independent uniform leaf l' of $\mathcal{C}$; in particular, we consider the associated sequences of subtrees $\mathcal{C}(u)$ and leaves $l(u)$, $l'(u)$, for $u\in U$. For all $n\in\mathbb{N}$, let 1_n denote the n-uple $(1,\ldots,1)$.

Let $n \in \mathbb{N}$, and consider a uniform leaf $\tilde{l}$ of $\mathcal{C}(1_n)$. Up to an appropriate rescaling, $\mathcal{C}(1_n)$ is a stable tree of index α and mass 1, on which the restriction of the decoration f and the distinguished point $l'(1_n)$ allow us to define a reconstructed distance $\delta_{\mathcal{C}(1_n)}$. As above, it can be seen from the construction of $\delta_{\mathcal{C}}$ that

$$\delta_{\mathcal{C}}(l^0, \tilde{l}) = m(1_n)^{1-1/\alpha} \delta_{\mathcal{C}(1_n)}(l^0, \tilde{l}). \tag{3.11}$$

Let

$$\Delta_{\mathcal{C}(1_n)} = \sup_{\tilde{l} \in \mathcal{D} \cap \mathcal{C}(1_n)} \delta_{\mathcal{C}(1_n)}(l^0, \tilde{l}).$$

The variable $\Delta_{\mathcal{C}(1_n)}$ is independent of $m(1_n)$ and has the same distribution as

$$\Delta := \sup_{\tilde{l} \in \mathcal{D}} \delta_{\mathcal{C}}(l^0, \tilde{l}).$$

(Note that we know from equation (3.2) that its law is that of the height of a stable tree of index α. Moreover, as suggested by the notation, $\Delta_{\mathcal{C}(1_n)}$ only depends on the rescaled subtree $\mathcal{C}(1_n)$, endowed with its decoration and its distinguished leaf). Equality 3.11 yields

$$\sup_{\tilde{l} \in \mathcal{D} \cap \mathcal{C}(1_n)} \delta_{\mathcal{C}}(l^0, \tilde{l}) \leq m(1_n)^{1-1/\alpha} \cdot \Delta_{\mathcal{C}(1_n)}.$$

We now show that there exists a subsequence of $(m(1_n)^{1-1/\alpha} \cdot \Delta_{\mathcal{C}(1_n)})_{n \in \mathbb{N}}$ converging to 0 almost surely, as $n \to \infty$. First note that $\mathbb{E}\left[m(1_n)^{1-1/\alpha}\right] = 2^{-n}$, so $m(1_n)$ converges to 0 in L^1. Since $(m(1_n))_{n \geq 0}$ is decreasing in n, this convergence also holds almost surely. Now

$$\mathbb{P}\left(m(1_n)^{1-1/\alpha} \cdot \Delta_{\mathcal{C}(1_n)} > \eta\right) = \mathbb{P}\left(\Delta > \frac{\eta}{m(1_n)^{1-1/\alpha}}\right) \xrightarrow[n \to \infty]{} 0.$$

Thus $m(1_n)^{1-1/\alpha} \cdot \Delta_{\mathcal{C}(1_n)}$ converges to 0 in probability, hence the desired convergence. As a consequence, there exists a (random) $n \in \mathbb{N}$ such that $m(1_n)^{1-1/\alpha} \cdot \Delta_{\mathcal{C}(1_n)} \leq \varepsilon/2$ almost surely. Taking $V_{l^0}(\varepsilon) = \mathcal{C}(1_n)$ yields the desired result. □

5.2 Correspondence between the leaves of $\mathcal{C}$ and the points of $\mathrm{Rec}(\mathcal{C})$

With the definition of the previous section, it is easy to see that the mapping

$$\begin{array}{rccc} \delta: & \mathcal{L}^\bullet(\mathcal{C})^2 & \to & \mathbb{R}_+ \\ & (x, y) & \mapsto & \delta(x, y) \end{array}$$

is continuous, and that it defines a pseudo-distance on $\mathcal{L}^\bullet(\mathcal{C})$.

Lemma 5.4. *For all $x \neq y \in \mathcal{L}^\bullet(\mathcal{C})$, we have $\delta(x, y) = 0$ if and only if $x = f_x(x \wedge y)$ and $y = f_y(x \wedge y)$.*

Proof. We begin with two remarks. First, for all $x, y \in \mathcal{L}^\bullet(\mathcal{C})$, we have

$$\delta(x, y) = \delta(x, f_x(b)) + \delta(f_y(b), y).$$

Indeed, the construction of δ shows that this is true if $x, y \in \mathcal{D}$, and it can be extended to $x, y \in \mathcal{L}^\bullet(\mathcal{C})$ using the continuity of δ and the density of $\mathcal{D}$ in $\mathcal{L}^\bullet(\mathcal{C})$. This shows the reverse implication. Second, almost surely, for all $l \neq l' \in \mathcal{D}$, we have $\delta(l(1,2), l'(1,2)) > 0$ (since $l(1,2), l'(1,2)$ are independent uniform leaves of the stable tree $\mathcal{C}(1,2)$).

Now choose $x \neq y \in \mathcal{L}^\bullet(\mathcal{C})$ such that, for example, $x \neq f_x(x \wedge y)$. Let $z = f_x(x \wedge y)$ and $b' = x \wedge z$. Since $\mu_{\mathcal{C}}(\mathcal{C}_{b'}^x) > 0$ and $\mu_{\mathcal{C}}(\mathcal{C}_b^y) > 0$, there exists $l \in (\mathcal{C}_{b'}^x \cap \mathcal{D}$ and $l' \in \mathcal{C}_b^y \cap \mathcal{D}$, and the associated leaves $l(1,2)$ and $l'(1,2)$ are respectively $f_z(b')$ and z. As a consequence, we have

$$\delta(x, y) = \delta(x, f_x(b')) + \delta(f_z(b'), z) + \delta(f_y(b), y) \geq \delta(f_z(b'), z) > 0.$$

□

Recall from Section 1 that the reconstructed tree was defined as the continuous random tree $(\mathrm{Rec}(\mathcal{C}), d_{\mathcal{R}}, \mu_{\mathcal{R}})$ in which $(\delta(l_i, l_j))_{i,j\geq 0}$ is the matrix of mutual distances between i.i.d. points. More precisely, there exists a mapping $r_{\mathcal{C}} : \mathcal{D} \cup \{l_0\} \to \mathrm{Rec}(\mathcal{C})$ such that for all i, j, we have $d_{\mathcal{R}}(r_{\mathcal{C}}(l_i), r_{\mathcal{C}}(l_j)) = \delta_{\mathcal{C}}(l_i, l_j)$, and

$$\mathrm{Rec}(\mathcal{C}) = \overline{\bigcup_{i\in\mathbb{N}} [\![r_{\mathcal{C}}(l_0), r_{\mathcal{C}}(l_i)]\!]}.$$

With this notation, for all $x \in \mathcal{L}^{\bullet}(\mathcal{C})$, there exists a unique point $r_{\mathcal{C}}(x) \in \mathrm{Rec}(\mathcal{C})$ such that

$$(d(r_{\mathcal{C}}(x), r_{\mathcal{C}}(l_i))_{i\geq 0} = (\delta_{\mathcal{C}}(x, l_i))_{i\geq 0}.$$

Indeed, the set

$$\bigcap_{\varepsilon>0} \overline{r_{\mathcal{C}}(V_x(\varepsilon))}$$

is non-empty (since $\mathrm{Rec}(\mathcal{C})$ is compact) and has diameter 0, so it is enough to define $r_{\mathcal{C}}(x)$ as its unique element. Therefore, we get the following property:

Proposition 5.5. *There exists a continuous map $r_{\mathcal{C}}$ from $\mathcal{L}^{\bullet}(\mathcal{C})$ into $\mathrm{Rec}(\mathcal{C})$ such that for all $x, y \in \mathcal{C}$, we have $d_{\mathcal{R}}(r_{\mathcal{C}}(x), r_{\mathcal{C}}(y)) = \delta_{\mathcal{C}}(x, y)$, and $r_{\mathcal{C}}(\mathcal{L}^{\bullet}(\mathcal{C}))$ is dense in $\mathrm{Rec}(\mathcal{C})$. Moreover, for all $x \neq y \in \mathcal{L}^{\bullet}(\mathcal{C})$, we have $r_{\mathcal{C}}(x) = r_{\mathcal{C}}(y)$ if and only if $x = f_x(x \wedge y)$ and $y = f_y(x \wedge y)$.*

Note that in particular, for all $b \in \mathcal{B}(\mathcal{C})$, the set $r_{\mathcal{C}}(f(b))$ has a unique element; in what follows, we call $r_{\mathcal{C}}(b)$ this point.

5.3 Back to the case where $\mathcal{C}$ is the cut-tree of a stable tree $\mathcal{T}$

Our main goal is to check that the reconstructed distance still coincides with the initial distance on $\mathcal{T}$, for all points for which it is well defined:

Proposition 5.6. *Almost surely, for all $X, Y \in \ell^{-1}(\mathcal{L}^{\bullet}(\mathcal{C}))$, we have $\delta(\ell_X, \ell_Y) = d_{\mathcal{T}}(X, Y)$.*

Note that for the points $\chi_{i,j}$, $(i, j) \in \mathbb{N} \times \mathcal{J}_\alpha$, the distance $d_{\mathcal{T}}$ is implicitly defined as

$$d_{\mathcal{T}}(\chi_{i,j}, Y) = d_{\mathcal{T}}(\chi_i, Y)$$

for all $Y \in \overline{\mathcal{T}}_{\mathcal{X}}$.

Let $(L_k)_{k\in\mathbb{N}}$ be a sequence of i.i.d. uniform leaves of $\mathcal{T}$, and $\mathcal{D}_{\mathcal{T}} = \{L_k : k \in \mathbb{N}\}$. Note that we almost surely have $\mathcal{D}_{\mathcal{T}} \subset \ell^{-1}(\mathcal{L}^{\bullet}(\mathcal{C}))$, and

$$\delta_{\mathcal{C}}(\ell_L, \ell_{L'}) = d_{\mathcal{T}}(L, L')$$

for all $L, L' \in \mathcal{D}_{\mathcal{T}}$ (see Proposition 1.7). The natural idea of the proof consists in using the continuity of δ to extend this identity. To this end, we also need a continuity result for ℓ, but it is clear that ℓ is not continuous for the topology τ_1 induced by the distance $d_{\mathcal{T}}$ on $\overline{\mathcal{T}}_{\mathcal{X}}$, since τ_1 does not separate the points $\chi_{i,j}$, $j \in \mathcal{J}_\alpha$ for a given $i \in \mathbb{N}$. This leads us to work with a topology τ_2 for which these points are separated, defined as follows:

— for all $x \in \mathcal{T} \setminus \mathcal{X}$, V is a neighbourhood of x for τ_2 if and only if it is a neighbourhood of x for τ_1,

— for all $(i, j) \in \mathbb{N} \times \mathcal{J}_\alpha$, V is a neighbourhood of $\chi_{i,j}$ for τ_2 if and only if there exists a neighbourhood V' of $\chi_{i,j}$ for τ_1 such that $V' \cap \overline{\mathcal{T}}(\chi_{i,j}, t_i) \subset V$.

(τ_2 can be seen as the topology generated by the union of the topologies on the subtrees $\overline{\mathcal{T}}(\chi_{i,j}, t_i)$, for $(i, j) \in \mathbb{N} \times \mathcal{J}_\alpha$.) Note that $\mathcal{D}_{\mathcal{T}}$ is still dense in $\mathcal{T}$ for the topology τ_2.

Lemma 5.7. *Almost surely, the application ℓ is continuous on $\ell^{-1}(\mathcal{L}^{\bullet}(\mathcal{C}))$ for the topology τ_2.*

Proof. It is clear from the construction of $\mathrm{Cut}(\mathcal{T})$ that for all $K \geq 0$, the quantity

$$D'_K := \sup_{e \in \mathcal{E}^*(K)} \sup_{L,L' \in \overline{\mathcal{T}}_e \cap \mathcal{D}_{\mathcal{T}}} d_{\mathcal{C}}(\ell_L, \ell_{L'})$$

has the same distribution as the quantities D_K and Δ_K introduced in the proof of Lemma 5.2. Thus D'_K converges to 0 almost surely as $K \to \infty$. Moreover, since for all $e \in \mathcal{E}_\alpha$, $\ell(\overline{\mathcal{T}}_e \cap \mathcal{D}_{\mathcal{T}})$ is a.s. dense in $\ell(\overline{\mathcal{T}}_e)$, we have

$$D'_K = \sup_{e \in \mathcal{E}^*(K)} \sup_{X,Y \in \overline{\mathcal{T}}_e} d_{\mathcal{C}}(\ell_X, \ell_Y),$$

hence the continuity of ℓ at every point of the set

$$\bigcap_{K \geq 0} \bigcup_{e \in \mathcal{E}^*(K)} \overline{\mathcal{T}}_e = \ell^{-1}(\mathcal{L}^\bullet(\mathcal{C})) \setminus \{r_e : e \in \mathcal{E}_\alpha\}.$$

(Indeed, for any element X of $\overline{\mathcal{T}}_e$, $e \in \mathcal{E}^*$, $\overline{\mathcal{T}}_e$ is a neighbourhood of X for the topology τ_2).

Now let $e \in \mathcal{E}_\alpha$, and consider the sequence $\mathcal{T}(u)$, $u \in U$ obtained by performing the construction of section 4 for $L = r_e$ and a uniform leaf L' in the (rescaled) stable tree $\overline{\mathcal{T}}_e$. An argument similar to that of Lemma 5.3 shows that almost surely, for n large enough, the diameter of $\mathcal{T}(1_n)$ is less than ε, hence the continuity of ℓ at point r_e. □

We can now complete the proof of the above result:

Proof of Proposition 5.6. Let $X, Y \in \ell^{-1}(\mathcal{L}^\bullet(\mathcal{C}))$. Consider two sequences $(X_n), (Y_n)$ in $\mathcal{D}_{\mathcal{T}} \cap \ell^{-1}(\mathcal{L}^\bullet(\mathcal{C}))$, respectively converging to X and Y for the topology τ_2. The continuity properties of ℓ and δ ensure that

$$\lim_{n \to \infty} \delta(\ell_{X_n}, \ell_{Y_n}) = \delta(\ell_X, \ell_Y).$$

Moreover, for all n, we have $X_n, Y_n \in \mathcal{D}_{\mathcal{T}}$, so

$$\delta(\ell_{X_n}, \ell_{Y_n}) = d_{\mathcal{T}}(X_n, Y_n).$$

Finally, since the topology τ_2 is finer than τ_1, we have

$$\lim_{n \to \infty} d_{\mathcal{T}}(X_n, Y_n) = d_{\mathcal{T}}(X, Y),$$

which yields the conclusion. □

In particular, we are interested in the reconstructed points corresponding to cut-points in $\mathcal{T}$. Recall that for all $i \in \mathbb{N}$, there exists a unique branching point b_i of $\mathcal{C}$ corresponding to the cut-point (χ_i, t_i), such that

$$f^{\mathcal{T}}(b_i) = \left\{\ell_{\chi_{i,j}} : j \in \mathcal{J}_\alpha\right\}.$$

Proposition 5.6 allows us to write, for any $j \in \mathcal{J}_\alpha$,

$$\begin{aligned}(d_{\mathcal{R}}(r_{\mathcal{C}}(b_i), r_{\mathcal{C}}(l_k)))_{k \geq 0} &= (d_{\mathcal{R}}(r_{\mathcal{C}}(\ell_{\chi_{i,j}}), r_{\mathcal{C}}(\ell_{L_k})))_{k \geq 0} \\ &= (\delta_{\mathcal{C}}(\ell_{\chi_{i,j}}, \ell_{L_k}))_{k \geq 0} \\ &= (d_{\mathcal{T}}(\chi_i, L_k))_{k \geq 0}.\end{aligned}$$

Thus the point $r_{\mathcal{C}}(b_i)$ of $\mathrm{Rec}(\mathcal{C})$ is "the same" as the cut-point χ_i of $\mathcal{T}$.

5.4 Fragmentation of the reconstructed tree

Here we explain how to recover a Poisson point process on $\mathrm{Rec}(\mathcal{C})$ (and in particular, the fragmentation times) such that the associated cut-tree is $\mathcal{C}$. We first assume that $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$ for a given stable tree $\mathcal{T}$. As above, we let $(L_k)_{k\in\mathbb{N}}$ be a sequence of i.i.d. uniform leaves of $\mathcal{T}$, $L_0 = \rho_{\mathcal{T}}$, and take $l_k = \ell_{L_k}$ for all $k \geq 0$.

Fix $k \geq 0$. For all $t \geq 0$, let

$$H_k(t) = \int_0^t \mu_{\mathcal{T}}(L_k, s)ds, \tag{3.12}$$

and recall that $\lim_{t\to\infty} H_k(t) = d_{\mathcal{C}}(\rho_{\mathcal{C}}, l_k)$. More precisely, for any atom (χ_i, t_i) of the Poisson point process used in the fragmentation of $\mathcal{T}$, there exists $k \geq 0$ such that the corresponding branching point b_i of $\mathcal{C}$ belongs to $[\![\rho_{\mathcal{C}}, l_k]\!]$, and we have

$$d_{\mathcal{C}}(\rho_{\mathcal{C}}, b_i) = H_k(t_i) \tag{3.13}$$

for any such k. We already know that we can recover the cut-point χ_i by taking $\chi_i = r_{\mathcal{C}}(b_i)$; to get the corresponding time t_i, we have to identify the inverse function of H_k. To this end, we use the fact that we know the correspondence between the masses of the connected components of the fragmented tree and the masses of the subtrees of $\mathcal{C}$. Indeed, for all $h \in [0, d_{\mathcal{C}}(\rho_{\mathcal{C}}, l_k)]$, letting $x_k(h)$ denote the unique point of $[\![\rho_{\mathcal{C}}, l_k]\!]$ such that $d_{\mathcal{C}}(\rho_{\mathcal{C}}, x_k(h)) = h$ and

$$M_k(h) = \mu_{\mathcal{C}}\left(\mathcal{C}^{l_k}_{x_k(h)}\right),$$

we have

$$M_k(H_k(t)) = \mu_{\mathcal{T}}(\overline{\mathcal{T}}(L_k, t)). \tag{3.14}$$

Equations (3.12) and (3.14) allow us to derive the following result.

Lemma 5.8. *For all $h \in [0, d_{\mathcal{C}}(\rho_{\mathcal{C}}, l_k)]$, the quantity*

$$T_k(h) = \int_0^h \frac{du}{M_k(u)}$$

is well-defined, and T_k is the inverse function of H_k.

Note that the definition of T_k does not use the fact that $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$. As a consequence, we have:

Proposition 5.9. *For all $b \in \mathcal{B}(\mathcal{C})$, the quantity $t_{\mathcal{C}}(b) := T_k(d_{\mathcal{C}}(\rho_{\mathcal{C}}, b))$ is independent of the choice of k such that $b \in [\![\rho_{\mathcal{C}}, l_k]\!]$, and of the choice of the i.i.d. sequence $(l_k)_{k\geq 0}$. With this notation:*

- *If $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$, then for all $i \in \mathbb{N}$, we have $r_{\mathcal{C}}(b_i) = \chi_i$ and $t_{\mathcal{C}}(b_i) = t_i$.*
- *In the general case, the points $(r_{\mathcal{C}}(b), t_{\mathcal{C}}(b))$, $b \in \mathcal{B}(\mathcal{C})$, have the same distribution as the atoms of a Poisson point process with intensity $\Lambda_\alpha(dt, dx)$ on $\mathbb{R}_+ \times \mathrm{Rec}(\mathcal{C})$. Moreover, for the fragmentation induced by these atoms, we have $\mathrm{Cut}(\mathrm{Rec}(\mathcal{C})) = \mathcal{C}$ almost surely.*

The case where $\mathcal{C} = \mathrm{Cut}(\mathcal{T})$ is a straightforward consequence of Lemma 5.8 and equation (3.13). For the general case, it is enough to note that since $\mathcal{C}$ has the same distribution as the cut-tree of $\mathcal{T}$, we have the joint equalities in distribution

$$\left(\mathrm{Rec}(\mathcal{C}), (r_{\mathcal{C}}(b), t_{\mathcal{C}}(b))_{b\in\mathcal{B}(\mathcal{C})}\right) \overset{(d)}{=} \left(\mathrm{Rec}(\mathrm{Cut}(\mathcal{T})), (r_{\mathrm{Cut}(\mathcal{T})}(b), t_{\mathrm{Cut}(\mathcal{T})}(b))_{b\in\mathcal{B}(\mathcal{C})}\right) = \left(\mathcal{T}, (\chi_i, t_i)_{i\in\mathbb{N}}\right),$$

and therefore,

$$(\mathcal{C}, \mathrm{Cut}(\mathrm{Rec}(\mathcal{C}))) \overset{(d)}{=} (\mathrm{Cut}(\mathcal{T}), \mathrm{Cut}(\mathcal{T})).$$

Chapitre 4

The UIPQ seen from a point at infinity along its geodesic ray

Les résultats de ce chapitre on été soumis pour publication dans Electronic Journal of Probability.

1 Introduction

Finite and infinite planar maps are a popular model for random geometry. While finite maps have been studied since the sixties, infinite models were only introduced a decade ago, with the works of Angel and Schramm [12, 10]. They were the first to define the uniform infinite planar triangulation, an infinite map which can be seen as the local limit (in distribution) of uniform finite triangulations. Krikun [42] then studied its counterpart, the uniform infinite planar quadrangulation (UIPQ), defined as the limit of uniform rooted finite quadrangulations as the number of faces goes to infinity. In this article, we study what the UIPQ looks like seen from a point "at infinity" on a geodesic ray originating from the root.

One of the main advantages of quadrangulations over other classes of planar maps is the existence of the so-called Cori-Vauquelin-Schaeffer bijection. This bijection, introduced in [24] and developed thoroughly in [56, 23], gives a correspondence between finite quadrangulations and well-labeled finite trees. It was in particular used by Chassaing and Durhuus [22] as a new approach to the UIPQ: they studied the infinite quadrangulation of the plane corresponding to an infinite positive labeled tree, and it was shown later by Ménard [52] that this quadrangulation has the same distribution as the one defined by Krikun.

Using another extension of the Cori-Vauquelin-Schaeffer bijection, Curien, Ménard and Miermont [27] recently showed that the UIPQ can also be obtained from a "uniform" infinite labeled tree, without the positivity constraint on the labels. This construction allowed them to prove new results on the UIPQ, and in particular to give a fine description of the geodesic arcs from a point to infinity. One of their main results states that all such geodesics are "trapped" between two distinguished geodesics, which have a simple description in terms of the corresponding labeled tree. Moreover, these two geodesics, called the maximal (or leftmost) and minimal (or rightmost) geodesics, are roughly similar, in the sense that they almost surely have an infinite number of common points.

Our main goal here is to study the local limit of $Q_\infty^{(k)}$ as $k \to \infty$, where $Q_\infty^{(k)}$ denotes the UIPQ re-rooted at a point at distance k from the root, on the leftmost geodesic. Our methods are again based on bijective correspondences between trees and quadrangulations. Specifically, we show that $Q_\infty^{(k)}$ converges in distribution to a limit quadrangulation $\overleftrightarrow{Q}_\infty$, which can be obtained by gluing together two quadrangulations of the half-plane with geodesic boundaries; we give explicit expressions for the distribution of the corresponding trees. Note that the laws of the quadrangulations of the half-plane we consider (corresponding to the parts of the UIPQ which are "on the left" and "on the right" of the leftmost geodesic ray) are orthogonal to the law of the uniform infinite quadrangulation of the half-plane (UIHPQ) which was studied in [11] and [26].

Finally, note that the scaling limit of the uniform infinite quadrangulation, the Brownian plane, which was introduced and studied by Curien and Le Gall [25], has a similar "uniqueness" property of infinite geodesic rays started from the root. We expect our result to have a natural analog in this context.

In the rest of this introduction, we give the necessary definitions to state our main results. In Section 1.1, we first recall classical definitions on quadrangulations and labeled trees; we also describe the construction of the UIPQ given in [27] and the "Schaeffer-type" correspondence it relies on. Section 1.2 gives more details on the UIPQ re-rooted at the k-th point on the leftmost infinite geodesic ray starting from the root. In particular, we explain why it is enough to study the local limit of the parts on each side of this geodesic. This leads us to extend the correspondence to a larger class of infinite labeled trees, which encode planar quadrangulations with a geodesic boundary (see Section 1.3). Finally, in Section 1.4, we state our main convergence results for these trees and the associated quadrangulations.

1.1 Well-labeled trees and associated quadrangulations

1.1.1 First definitions on finite and infinite planar maps

A finite planar map is a proper embedding of a finite connected graph, possibly with multiple edges or loops, into the two-dimensional sphere (or more rigorously, the equivalence class of such a graph, modulo orientation-preserving homeomorphisms).

We first introduce some notation for such a map $\mathfrak{m}$. Let $V(\mathfrak{m})$, $E(\mathfrak{m})$ and $\overrightarrow{E}(\mathfrak{m})$ denote the sets of the vertices, edges and oriented edges of $\mathfrak{m}$, respectively. The faces of $\mathfrak{m}$ are the connected components of the complement of $E(\mathfrak{m})$. We say that a face is incident to $e \in \overrightarrow{E}(\mathfrak{m})$ if it is the face on the left of e. The degree of a face is the number of edges it is incident to. A corner of $\mathfrak{m}$ is an angular sector between two edges of $\mathfrak{m}$. Note that there is a bijective correspondence between the corners of $\mathfrak{m}$ and its oriented edges; we say that a corner is incident to $e \in \overrightarrow{E}(\mathfrak{m})$ if it is the corner on the left of e, next to its origin.

We say that a finite planar map is rooted if it comes with a distinguished oriented edge, called the root edge; the origin vertex of the root is called the root vertex, and the face which is incident to the root is called the root face. A planar map is a quadrangulation if all faces have degree 4, and a tree if it has only one face. A quadrangulation with a boundary is a planar map with a distinguished face called the external face, such that the boundary of the external face is simple and all other faces have degree 4. We let $\mathcal{Q}_f$, $\mathcal{Q}_{f,b}$ and $\mathcal{T}_f$ respectively denote the sets of finite quadrangulations, quadrangulations with a boundary and trees.

Let us now define the local limit topology on these sets. For any rooted map $\mathfrak{m}$, let $B_{\mathfrak{m}}(r)$ denote the ball of radius r in $\mathfrak{m}$, centered at the root-vertex (i.e. the planar map defined by the edges of $\mathfrak{m}$ whose extremities are both at distance at most r from the root-vertex, for the graph-distance on $\mathfrak{m}$). For all finite planar maps $\mathfrak{m}$, $\mathfrak{m}'$, we let

$$D(\mathfrak{m}, \mathfrak{m}') = (1 + \sup\{r \geq 0 : B_{\mathfrak{m}}(r) = B_{\mathfrak{m}'}(r)\})^{-1}.$$

The local topology is the topology associated to this distance. Let $\mathcal{Q}$, $\mathcal{Q}_b$ and $\mathcal{T}$ denote the completions of $\mathcal{Q}_f$, $\mathcal{Q}_{f,b}$ and $\mathcal{T}_f$ for this topology. The elements of $\mathcal{Q}_\infty := \mathcal{Q} \setminus \mathcal{Q}_f$ (resp. $\mathcal{T}_\infty := \mathcal{T} \setminus \mathcal{T}_f$) are *infinite* planar quadrangulations (resp. trees). All the notations introduced above for finite planar maps have natural extensions to the above sets. We let $\mathcal{Q}_{\infty,\infty}$ denote the set of the quadrangulations with an infinite boundary, i.e. the elements of $\mathcal{Q}_b$ which are defined as limits of sequences of maps in $\mathcal{Q}_{f,b}$ whose external faces have degrees going to infinity.

Any element Q of $\mathcal{Q}_\infty$ or $\mathcal{Q}_{\infty,\infty}$ can be seen as a gluing of quadrangles which defines an orientable, connected, separable surface, with a boundary in the second case. See [27, Appendix] for details. We are interested in two cases:

— If the corresponding surface is homeomorphic to $S = \mathbb{R}^2$, we say that Q is an infinite quadrangulation of the plane.

— If the corresponding surface is homeomorphic to $S = \mathbb{R} \times \mathbb{R}_+$, we say that Q is an infinite quadrangulation of the half-plane.

In both of these cases, Q can be drawn onto S in such a way that every face is bounded, every compact subset of S intersects only finitely many edges of Q, and in the second case, the union of the boundary edges is $\mathbb{R} \times \{0\}$. By convention, if the root edge belongs to this boundary and is oriented from left to right, we say that Q is a quadrangulation of the upper half-plane, and if it is oriented from right to left, we say that Q is a quadrangulation of the lower half-plane. We let $\mathbf{Q}$ denote the set of the quadrangulations of the plane, and $\overleftarrow{\mathbf{Q}}$ (resp. $\overrightarrow{\mathbf{Q}}$) denote the set of the infinite quadrangulations Q of the upper half-plane (resp. lower half-plane) such that the boundary of Q is a geodesic path in Q.

As explained in [27, Appendix], an element Q of $\mathcal{Q}_\infty$ is a quadrangulation of the plane if and only if it has exactly one *end* - which means, in terms of maps, that for all $r \in \mathbb{N}$, the map $Q \setminus B_Q(r)$ has exactly one infinite connected component. For an element Q of $\mathcal{Q}_{\infty,\infty}$, one can check that Q is a quadrangulation of the half-plane if and only if the same condition holds. Indeed, the infinite quadrangulation obtained by gluing a copy of the lattice $\mathbb{Z} \times \mathbb{Z}_-$ along the boundary also has one end (the number of ends can only decrease when we perform this operation), so it is a quadrangulation of the plane.

In what follows, the trees and quadrangulations we consider will be elements of $\mathbf{T} := \mathcal{T}$, $\mathbf{Q}$, $\overleftarrow{\mathbf{Q}}$ and $\overrightarrow{\mathbf{Q}}$. The uniform infinite quadrangulation (UIPQ) is a random variable in $\mathbf{Q}$ whose distribution is the limit of the uniform distribution on planar quadrangulations with n faces, as $n \to \infty$.

1.1.2 Well-labeled trees

We say that (T, l) is a well-labeled (plane, rooted) tree if T is an element of $\mathbf{T}$ and l is a mapping from $V(T)$ into $\mathbb{Z}$ such that $|l(\mathrm{u}) - l(\mathrm{v})| \leq 1$ for every pair of neighbouring vertices u, v. Let $\mathbb{T}$ be the set of such trees. More precisely, for all $x \in \mathbb{Z}$ and $n \geq 0$, let $\mathbb{T}_n(x)$ be the set of well-labeled plane rooted trees with n edges and root-label x, and

$$\mathbb{T}_n = \bigcup_{x \in \mathbb{Z}} \mathbb{T}_n(x).$$

Similarly, for all $x \in \mathbb{Z}$, let $\mathbb{T}_\infty(x)$ denote the set of infinite well-labeled plane rooted trees with root-label x, and

$$\mathbb{T}_\infty = \bigcup_{x \in \mathbb{Z}} \mathbb{T}_\infty(x).$$

We thus have

$$\mathbb{T} = \bigcup_{n \in \mathbb{N} \cup \{0, \infty\}} \mathbb{T}_n = \bigcup_{n \in \mathbb{N} \cup \{0, \infty\}} \bigcup_{x \in \mathbb{Z}} \mathbb{T}_n(x).$$

For any (infinite) plane rooted tree T, we say that $(\mathrm{u}_i)_{i \geq 0}$ is a spine in T if u_0 is the root of T and if for all $i \geq 0$, u_i is the parent of u_{i+1}. We let $\mathbf{S}$ be the set of all plane rooted trees having exactly one spine, and consider the corresponding sets of labeled trees:

$$\mathbb{S}(x) = \{(T, l) \in \mathbb{T}_\infty(x) : T \in \mathbf{S}\} \qquad \forall x \in \mathbb{Z},$$
$$\mathbb{S} = \{(T, l) \in \mathbb{T}_\infty : T \in \mathbf{S}\}.$$

For every $T \in \mathbf{S}$, we let $(\mathfrak{s}_i(T))_{i \geq 0}$ be the spine of T. Any vertex $\mathfrak{s}_i(T)$ has a subtree "to its left" and a subtree "to its right" in T, which we denote by $L_i(T)$ and $R_i(T)$ respectively. To give a formal definition of these subtrees, we consider two orders on $V(T)$: the depth-first order, denoted by $<$, and the partial order $\prec$ induced by the genealogy, defined for all $\mathrm{u}, \mathrm{v} \in V(T)$ by $\mathrm{u} \prec \mathrm{v}$ if u is an ancestor of v in T. With this notation:

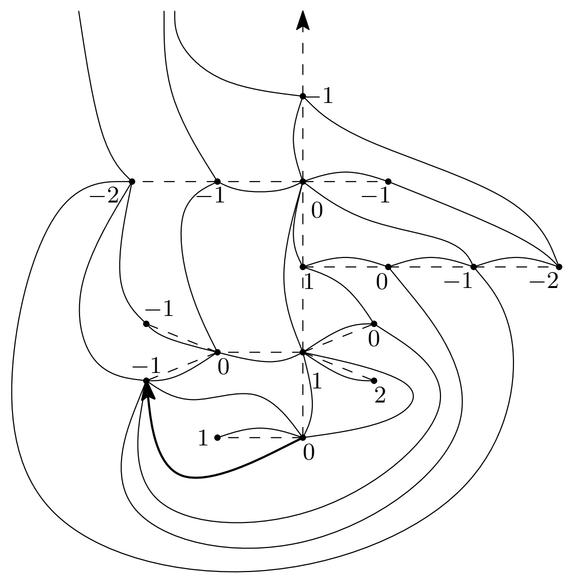

Figure 4.1 – The quadrangulation $\Phi(\theta)$ obtained by applying the Schaeffer correspondence to a labeled tree θ. The edges of θ are represented by dashed lines.

— $L_i(T)$ is the subtree of T containing the vertices v such that $\mathfrak{s}_i \leq \mathrm{v} < \mathfrak{s}_{i+1}$.
— $R_i(T)$ is the subtree of T containing $\mathfrak{s}_i$ and the vertices v such that $\mathfrak{s}_{i+1} < \mathrm{v}$ and $\mathfrak{s}_{i+1} \not\prec \mathrm{v}$.
We also use the natural extensions of these notations to well-labeled trees.

1.1.3 The Schaeffer correspondence between infinite trees and quadrangulations

In this section, we recall the definition of the Schaeffer correspondence used in [27], which matches infinite well-labeled trees with infinite quadrangulations of the plane.

For all $x \in \mathbb{Z}$, let

$$\mathbb{S}^*(x) = \left\{ (T, l) \in \mathbb{S}(x) : \inf_{i \geq 0} l(\mathfrak{s}_i(T)) = -\infty \right\}.$$

We fix $\theta = (T, l) \in \mathbb{S}^*(0)$. Let c_n, $n \in \mathbb{Z}$ denote the corners of T, taken in the clockwise order, with c_0 the root-corner. For all n, we say that the label of c_n is the label of the vertex which is incident to c_n, and we define the successor $\sigma_\theta(\mathrm{c}_n)$ of c_n as the first corner among $\mathrm{c}_{n+1}, \mathrm{c}_{n+2}, \ldots$ such that

$$l(\sigma_\theta(\mathrm{c}_n)) = l(\mathrm{c}_n) - 1.$$

We now let $\Phi(\theta)$ denote the graph whose set of vertices is $V(T)$, whose edges are the pairs $\{\mathrm{c}, \sigma_\theta(\mathrm{c})\}$ for all corners c of T, and whose root-edge is $(\mathrm{c}_0, \sigma_\theta(\mathrm{c}_0))$. Figure 4.1 gives an example of this construction. Note that $\Phi(\theta)$ can be embedded naturally in the plane, by considering a specific embedding of T and drawing arcs between every corner and its successor in a non-crossing way. Moreover, Proposition 2 of [27] shows that for all $\theta \in \mathbb{S}^*(0)$, $\Phi(\theta)$ is an infinite quadrangulation of the plane.

For a technical reason, we extend this definition to trees $\theta \in \mathbb{S}^*(1)$ by keeping the same vertices and edges, and choosing $(\sigma_\theta(\mathrm{c}_0), \sigma_\theta(\sigma_\theta(\mathrm{c}_0)))$ as the root. (Thus the root edge of $\Phi(\theta)$ always goes from vertices with labels 0 and -1 in θ.) For all $\theta \in \mathbb{S}^*(1)$, we still have $\Phi(\theta) \in \mathbf{Q}$.

1.1.4 Uniform infinite labeled tree and quadrangulation

For all $x \in \mathbb{Z}$, let $\rho_{(x)}$ be the law of a Galton–Watson tree with offspring distribution Geom(1/2), such that the root has label x and, for any vertex v other than the root, the label of v is uniform in $\{\ell-1, \ell, \ell+1\}$, with ℓ the label of its parent. The uniform infinite labeled tree is the random variable $\theta_\infty = (T_\infty, l_\infty) \in \mathbb{S}(0)$ whose distribution is characterized by the following properties:

— the process of the spine-labels $(S_i(\theta_\infty))_{i\geq 0} := (l_\infty(\mathfrak{s}_i(T_\infty)))_{i\geq 0}$ is a random walk with independent uniform steps in $\{-1, 0, 1\}$,

— conditionally on $(S_i(\theta_\infty))_{i\geq 0}$, the trees $L_i(\theta_\infty)$ and $R_i(\theta_\infty)$ are independent labeled trees distributed according to $\rho_{(S_i)}$.

For all $n \in \mathbb{N}$, we also let $\theta_n = (T_n, l_n)$ be a uniform random element of $\mathbb{T}_n(0)$. It is known that θ_n converges to θ_∞ for the local limit topology, as $n \to \infty$ (as noted in [27], it is a consequence of [41, Lemma 1.14]). Note that we have $\theta_\infty \in \mathbb{S}^*(0)$ almost surely, and let $Q_\infty := \Phi(\theta_\infty)$.

It was shown in [27] that the UIPQ can be seen as the random quadrangulation $\widetilde{Q}_\infty$ equal to Q_∞ with probability 1/2, and to the quadrangulation obtained by reversing the root edge of Q_∞ with probability 1/2.

1.2 Re-rooting the UIPQ at the k-th point on the leftmost geodesic ray

Let us first clarify what we mean by the leftmost geodesic originating form the root in the UIPQ. It is known from [27] that for all vertices u, v of $\widetilde{Q}_\infty$, the quantity

$$\lim_{\mathrm{w}\to\infty} \left(d_{\widetilde{Q}_\infty}(\mathrm{u}, \mathrm{w}) - d_{\widetilde{Q}_\infty}(\mathrm{v}, \mathrm{w}) \right)$$

is well defined (in the sense that the difference of those distances is the same except for a finite number of vertices w), and equal to the difference of the labels of u and v in the corresponding tree. As a consequence, letting e denote the root edge of Q_∞, with e^- its origin and e^+ its other extremity, we have

$$\lim_{\mathrm{w}\to\infty} \left(d_{\widetilde{Q}_\infty}(e^-, \mathrm{w}) - d_{\widetilde{Q}_\infty}(e^+, \mathrm{w}) \right) = 1.$$

In other words, the extremity of the root edge of $\widetilde{Q}_\infty$ which is “closest to infinity” is well defined, and equal to e^+. Therefore, it is natural to say that the leftmost geodesic ray started from the root in $\widetilde{Q}_\infty$ is the unique path $\gamma_L = (\gamma_L(i))_{i\geq 0}$ such that $\gamma_L(0) = e^-$, $\gamma_L(1) = e^+$ and for all $i \geq 1$, $\gamma_L(i+1)$ is the first neighbour of $\gamma_L(i)$ after $\gamma_L(i-1)$ (in the clockwise order) such that

$$\lim_{\mathrm{w}\to\infty} \left(d_{\widetilde{Q}_\infty}(\gamma_L(i), \mathrm{w}) - d_{\widetilde{Q}_\infty}(\gamma_L(i+1), \mathrm{w}) \right) = 1.$$

Note that the definition of the leftmost geodesic ray does not depend on whether the root edge of $\widetilde{Q}_\infty$ has the same orientation as that of Q_∞ or not, so it is sufficient to work with Q_∞ in the rest of the article.

The leftmost geodesic also has a natural definition in terms of the tree θ_∞. For all $k \geq 0$, let e_k be the k-th corner on the chain of the iterated successors of e_0, where e_0 is the root corner of θ_∞. Equivalently, e_k can be seen as the first corner with label $-k$ after the root, in the clockwise order. We use the same notation for the corresponding vertex in Q_∞. The path $\gamma_{\max} := (\mathrm{e}_k)_{k\geq 0}$ is a geodesic ray in Q_∞, called the maximal geodesic in [27], and equal to γ_L.

Curien, Ménard and Miermont proved in [27] that all other geodesic rays from e_0 to infinity are essentially similar to $\gamma_{\max}$: almost surely, there exists an infinite sequence of distinct vertices of Q_∞ such that every geodesic ray from e_0 to infinity passes through all these vertices. Our main goal is to study the local limit of $Q_\infty^{(k)}$ as $k \to \infty$, where $Q_\infty^{(k)}$ denotes the quadrangulation Q_∞ re-rooted at $(\mathrm{e}_k, \mathrm{e}_{k+1})$.

More precisely, we will study what the quadrangulation looks like on the left and on the right of the geodesic ray $\gamma_{\max}$. This leads us to introduce the “split” quadrangulation $\mathrm{Sp}(Q_\infty)$ obtained by “cutting” Q_∞ along $\gamma_{\max}$; formally, $\mathrm{Sp}(Q_\infty)$ is an infinite quadrangulation of the (lower) half-plane whose boundary is formed by the edges $(\mathrm{e}_k, \mathrm{e}_{k+1})$ on the left of e_0, and by copies $(\mathrm{e}'_k, \mathrm{e}'_{k+1})$ of these edges on the right of e_0. This construction is illustrated in Figure 4.2.

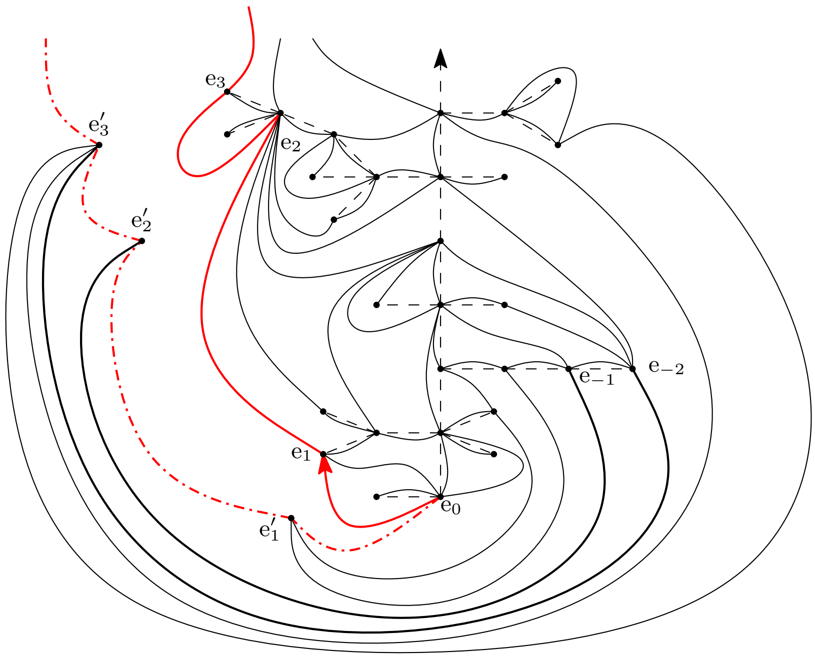

Figure 4.2 – The "split" quadrangulation $\mathrm{Sp}(Q_\infty)$ obtained from θ_∞. The edges of the underlying tree θ_∞ are represented in dashed lines, and the geodesic ray $\gamma_{\max}$ is represented in red. The labels are omitted to keep the figure readable.

For all $k \geq 0$, we let $\overrightarrow{Q}_\infty^{(k)}$ denote the quadrangulation having the same vertices and edges as $\mathrm{Sp}(Q_\infty)$, with root $(\mathrm{e}_k, \mathrm{e}_{k+1})$, and $\overleftarrow{Q}_\infty^{(k)}$ denote the quadrangulation having the same vertices and edges as $\mathrm{Sp}(Q_\infty)$, with root $(\mathrm{e}'_k, \mathrm{e}'_{k+1})$. Thus, since $(\mathrm{e}_k)_{k\geq 0}$ and $(\mathrm{e}'_k)_{k\geq 0}$ are geodesics in $\overrightarrow{Q}_\infty^{(k)}$ and $\overleftarrow{Q}_\infty^{(k)}$, we have the following property:

Lemma 1.1. *For all $r \leq k$, the ball of radius r in $Q_\infty^{(k)}$ is the same as the union of the balls of radius r in $\overrightarrow{Q}_\infty^{(k)}$ and $\overleftarrow{Q}_\infty^{(k)}$.*

The main idea now consists in studying the limit of the trees encoding $\overrightarrow{Q}_\infty^{(k)}$ and $\overleftarrow{Q}_\infty^{(k)}$, and then going back to the associated quadrangulations.

To this end, for all $k \in \mathbb{N}$, we introduce the tree $\theta_\infty^{(k)} = (T_\infty^{(k)}, l_\infty^{(k)})$, where $T_\infty^{(k)}$ is the tree T_∞ re-rooted at e_k, and $l_\infty^{(k)} := l_\infty + k$. Note that the vertices e'_k, $k \geq 1$, contrary to the e_k, do not correspond to corners of the tree θ_∞. Therefore, for all $k \in \mathbb{N}$, we let e_{-k+1} denote the last corner of θ_∞ before the root (still in the clockwise order) such that $\sigma_{\theta_\infty}(\mathrm{e}_{-k+1}) = \mathrm{e}_k$. Equivalently, e_{-k+1} can be seen as the last corner with label $-k+1$ before the root (hence the choice of the index). Now, for all $k \in \mathbb{N}$, we let $\theta_\infty^{(-k+1)} = (T_\infty^{(-k+1)}, l_\infty^{(-k+1)})$, where $T_\infty^{(-k+1)}$ is the tree T_∞ re-rooted at e_{-k+1}, and $l_\infty^{(-k+1)} := l_\infty + k$. With this notation, for all $k \in \mathbb{N}$, we have $\theta_\infty^{(k)} \in \mathbb{S}^*(0)$, $\theta_\infty^{(-k+1)} \in \mathbb{S}^*(1)$, $\Phi(\theta_\infty^{(k)}) = \overrightarrow{Q}_\infty^{(k)}$ and $\Phi(\theta_\infty^{(-k+1)}) = \overleftarrow{Q}_\infty^{(k)}$; but more importantly, we will show in Section 4 that the local limits of $\overrightarrow{Q}_\infty^{(k)}$ and $\overleftarrow{Q}_\infty^{(k)}$ can be determined using the local limits of $\theta_\infty^{(k)}$ and $\theta_\infty^{(-k+1)}$.

Intuitively, one can anticipate that the local limit of $\theta_\infty^{(k)}$ will be a tree in which the right-hand side only has positive labels, and the local limit of $\theta_\infty^{(-k+1)}$ will be a tree in which the left-hand side only has labels greater than 1. This leads us to extend the domain of Φ to such trees.

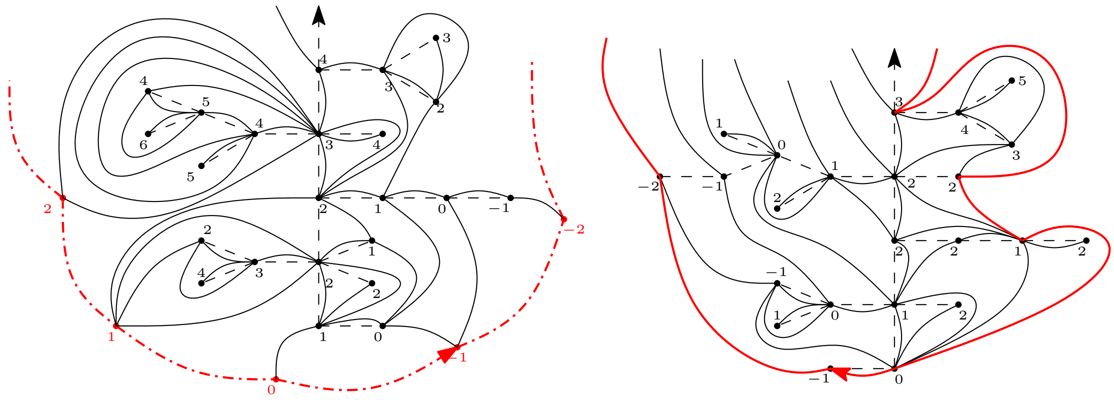

Figure 4.3 – Examples of quadrangulations $\Phi(\theta)$ obtained for $\theta \in \overleftarrow{\mathbb{S}}$ (on the left-hand side) and $\theta \in \overrightarrow{\mathbb{S}}$ (on the right-hand side).

1.3 Extending the Schaeffer correspondence

Consider the following subsets of $\mathbb{S}$:

$$\overrightarrow{\mathbb{S}} = \left\{ (T,l) \in \mathbb{S}(0) : \min_{n \leq -1} l(\mathrm{c}_n(T)) = 1, \lim_{n \to -\infty} l(\mathrm{c}_n(T)) = +\infty \text{ and } \inf_{n \geq 0} l(\mathrm{c}_n(T)) = -\infty \right\}$$

$$\overleftarrow{\mathbb{S}} = \left\{ (T,l) \in \mathbb{S}(1) : \min_{n \geq 1} l(\mathrm{c}_n(T)) = 2, \lim_{n \to +\infty} l(\mathrm{c}_n(T)) = +\infty \text{ and } \inf_{n \leq 0} l(\mathrm{c}_n(T)) = -\infty \right\}.$$

Here, we show that "Schaeffer-type" constructions yield natural associations between the trees in these sets and quadrangulations of the lower and upper half-planes. Examples of quadrangulations obtained this way are given on Figure 4.3.

In the case where $\theta \in \overrightarrow{\mathbb{S}}$, the construction is exactly the same as for $\theta \in \mathbb{S}^*(0)$: for all n, we define the successor $\sigma_\theta(\mathrm{c}_n)$ of c_n as the first corner among $\mathrm{c}_{n+1}, \mathrm{c}_{n+2}, \ldots$ such that

$$l(\sigma_\theta(\mathrm{c}_n)) = l(\mathrm{c}_n) - 1,$$

and we let $\Phi(\theta)$ denote the graph whose set of vertices is $V(T)$, whose edges are the pairs $\{\mathrm{c}, \sigma_\theta(\mathrm{c})\}$ for all corners c of T, and whose root-edge is $(\mathrm{c}_0, \sigma_\theta(\mathrm{c}_0))$.

Now, consider the case where $\theta \in \overleftarrow{\mathbb{S}}$. If we use the above construction, then for example, for all i, the last corner with label i has no successor. We therefore add a "shuttle" Λ, i.e. a line of new points λ_i, $i \in \mathbb{Z}$ on which the corners with no successor will be attached. More precisely, for all n, the successor of c_n is defined as

$$\sigma_\theta(\mathrm{c}_n) = \begin{cases} \mathrm{c}_{n'} & \text{for the smallest } n' \geq n \text{ such that } l(\mathrm{c}_{n'}) = l(\mathrm{c}_n) - 1, \text{ if it exists,} \\ \lambda_{l(\mathrm{c}_n)-1} & \text{otherwise,} \end{cases}$$

and we extend this notation to the points of Λ by letting $\sigma_\theta(\lambda_i) = \lambda_{i-1}$ for all $i \in \mathbb{Z}$. We let $\Phi(\theta)$ be the graph whose set of vertices is $V(T) \sqcup \Lambda$, whose edges are the pairs $\{\mathrm{c}, \sigma_\theta(\mathrm{c})\}$ for all corners c of T, and the pairs $\{\lambda_i, \lambda_{i-1}\}$ for all $i \in \mathbb{Z}$, and whose root-edge is $(\lambda_0, \lambda_{-1})$. (Note that the rooting convention is consistent with the one we used to define Φ on $\mathbb{S}^*(1)$.)

Lemma 1.2. *We have the following properties:*
- *If $\theta \in \overrightarrow{\mathbb{S}}$, then $\Phi(\theta) \in \overrightarrow{\boldsymbol{Q}}$.*
- *If $\theta \in \overleftarrow{\mathbb{S}}$, then $\Phi(\theta) \in \overleftarrow{\boldsymbol{Q}}$.*

Proof. In both cases, it is clear that the graph $\Phi(\theta)$ has a natural embedding into the plane, and the conditions on $\liminf_{n \to -\infty} l(\mathrm{c}_n)$ ensure that every corner is the successor of a finite number of other corners. Thus every vertex of $\Phi(\theta)$ has finite degree: $\Phi(\theta)$ is an infinite planar map.

As in Schaeffer's usual construction, a simple case study shows that for every corner c of θ:

— The face which is on the right of $(\mathrm{c}, \sigma_\theta(\mathrm{c}))$ is a quadrangle.

— If there exists a corner $\mathrm{c}' < \mathrm{c}$ such that $l(\mathrm{c}') = l(\mathrm{c})$, then the face which is on the left of $(\mathrm{c}, \sigma_\theta(\mathrm{c}))$ is a quadrangle. If $\theta \in \overleftarrow{\mathbb{S}}$, this is always true. If $\theta \in \overrightarrow{\mathbb{S}}$, then the only corners for which it is not true are the c_{n_i}, $i \in \mathbb{Z}$, with $n_i = \min\{n \in \mathbb{Z} : l(\mathrm{c}_n) = i\}$. For all i, we have $\sigma(\mathrm{c}_{n_i}) = \mathrm{c}_{n_{i-1}}$, and the face which is on the left of $(\mathrm{c}_{n_i}, \mathrm{c}_{n_{i-1}})$ is the root face of $\Phi(\theta)$.

For $\theta \in \overleftarrow{\mathbb{S}}$, we also have to study the faces which are on the left and on the right of the edges $(\lambda_i, \lambda_{i-1})$: we easily see that the first one is always a quadrangle, and that the second one is the same for all i. Thus:

— For all $\theta \in \overrightarrow{\mathbb{S}}$, we have $\Phi(\theta) \in \mathcal{Q}_{\infty,\infty}$.

— For all $\theta \in \overleftarrow{\mathbb{S}}$, letting $\overline{\Phi(\theta)}$ denote the map obtained by reversing the root edge of $\Phi(\theta)$, we have $\overline{\Phi(\theta)} \in \mathcal{Q}_{\infty,\infty}$.

Note that the construction ensures that the classical bound

$$d_{\Phi(\theta)}(\mathrm{u}, \mathrm{v}) \geq |l(\mathrm{u}, \mathrm{v})| \tag{4.1}$$

still holds. As a consequence, in both cases, the boundary is a geodesic path. Moreover, the fact that θ has exactly one spine implies, by construction, that $\Phi(\theta)$ is one-ended. □

Note that for $\theta \in \overrightarrow{\mathbb{S}}$, for all $i < i'$, the path $(\lambda_j)_{i \leq j \leq i'}$ is the unique geodesic between λ_i and $\lambda_{i'}$. Indeed, all neighbours of $\lambda_{i'}$ different of $\lambda_{i'-1}$ have labels equal to $i+1$, so they are at distance (at least) $i' - i + 1$ from λ_i. In other words, the boundary is the *unique* geodesic path between vertices of Λ.

1.4 Main results

The first part of our work is the identification of the limit of the joint distribution of $(\theta_\infty^{(k)}, \theta_\infty^{(-k+1)})$ as $k \to \infty$. We begin by using the convergence of θ_n towards θ_∞ to give an explicit description of this joint distribution.

To give a more precise idea of these results, we adapt the notation of Section 1.2 to possibly finite trees. For all $\theta = (T, l) \in \mathbb{T}(0)$ and $k \geq 0$ such that $\min_{V(T)} l \leq -k$, let $\mathrm{e}_k(\theta)$ be the first corner having label $-k$ after the root, in clockwise order, $\mathrm{e}_{-k}(\theta)$ be the last corner having label $-k$ before the root, and $\mathrm{v}_k(\theta)$ be the most recent common ancestor of $\mathrm{e}_k(\theta)$ and $\mathrm{e}_{-k+1}(\theta)$. Note that for $k = 0$, this is well defined since $\mathrm{e}_0(\theta) = \mathrm{e}_{-0}(\theta)$. Finally, we define the finite analogs of $\theta_\infty^{(k)}$ and $\theta_\infty^{(-k+1)}$: conditionally on $\min_{V(T_n)} l_n \leq -|k|$, for all $n, k \in \mathbb{N}$, we let

— $\theta_n^{(k)} = (T_n^{(k)}, l_n^{(k)})$, where $T_n^{(k)}$ is the tree T_n re-rooted at $\mathrm{e}_k(\theta_n)$, and $l_n^{(k)} = l_n + k$,

— $\theta_n^{(-k+1)} = (T_n^{(-k+1)}, l_n^{(-k+1)})$, where $T_n^{(-k+1)}$ is the tree T_n re-rooted at $\mathrm{e}_{-k+1}(\theta_n)$, and $l_n^{(-k+1)} = l_n + k$.

It is easy to see that:

Lemma 1.3. *We have the joint convergence in distribution*

$$(\theta_n^{(k)}, \theta_n^{(-k+1)}) \xrightarrow[n\to\infty]{} (\theta_\infty^{(k)}, \theta_\infty^{(-k+1)}) \tag{4.2}$$

for the local limit topology.

Indeed, the operations which consist in re-rooting a tree $\theta \in \mathbb{T}$ at $\mathrm{e}_k(\theta)$ and $\mathrm{e}_{-k+1}(\theta)$ are both continuous for the local limit topology on $\mathbb{S}^*(0)$. Since θ_∞ belongs to this set, this yields the conclusion. This lemma will allow us to give an explicit description of the joint distribution of $\theta_\infty^{(k)}$ and $\theta_\infty^{(-k+1)}$ (see Proposition 2.1 for the distribution of $\theta_\infty^{(k)}$ alone, and Corollary 3.2 for the joint distribution).

We use these results to prove the convergence theorem below. Recall that $\rho_{(x)}$ denotes the distribution of a Galton–Watson tree with Geom(1/2) offspring distribution and “uniform” labels, with root label x. If x is positive, we let $\rho^+_{(x)}$ denote the same distribution, conditioned to have only positive labels. We also introduce a Markov chain $\tilde{X}$ taking values in $\mathbb{N}$, with transition probabilities

$$\begin{aligned} p_x &:= \mathbb{P}(\tilde{X}_1 = x+1 | \tilde{X}_0 = x) = \frac{(x+4)(2x+5)}{3(x+2)(2x+3)} \\ r_x &:= \mathbb{P}(\tilde{X}_1 = x | \tilde{X}_0 = x) = \frac{x(x+3)}{3(x+1)(x+2)} \\ q_x &:= \mathbb{P}(\tilde{X}_1 = x-1 | \tilde{X}_0 = x) = \frac{(x-1)(2x+1)}{3(x+1)(2x+3)}. \end{aligned}$$

Note that $\tilde{X}$ can be seen as a discrete version of a seven-dimensional Bessel process. Indeed, a theorem of Lamperti [43] shows that, under some easily checked conditions, the rescaled process $((1/\sqrt{n}) \cdot \tilde{X}_{\lfloor nt \rfloor})_{t \geq 0}$ converges in distribution to a diffusion process with generator

$$L = \frac{\alpha}{x} \frac{d}{dx} + \frac{\beta}{2} \frac{d^2}{dx^2},$$

where

$$\alpha = \lim_{x \to \infty} x \mathbb{E}[\tilde{X}_1 - \tilde{X}_0 | \tilde{X}_0 = x] = 2$$

and

$$\beta = \lim_{x \to \infty} \mathbb{E}[(\tilde{X}_1 - \tilde{X}_0)^2 | \tilde{X}_0 = x] = \frac{2}{3},$$

hence in our case

$$L = \frac{2}{3} \left(\frac{3}{x} \frac{d}{dx} + \frac{1}{2} \frac{d^2}{dx^2} \right).$$

Thus, $((1/\sqrt{n}) \cdot \tilde{X}_{\lfloor nt \rfloor})_{t \geq 0}$ converges to $(Z_{2t/3})_{t \geq 0}$, where Z denotes a Bessel(7) process started from 0.

Theorem 1.4. *We have the joint convergence in distribution*

$$(\theta^{(k)}_\infty, \theta^{(-k+1)}_\infty) \xrightarrow[k \to \infty]{} (\overrightarrow{\theta_\infty}, \overleftarrow{\theta_\infty}) \tag{4.3}$$

for the local topology, where $\overrightarrow{\theta_\infty} = (\overrightarrow{T_\infty}, \overrightarrow{l_\infty})$ *and* $\overleftarrow{\theta_\infty} = (\overleftarrow{T_\infty}, \overleftarrow{l_\infty})$ *are independent random variables in* $\mathbb{S}(0)$ *and* $\mathbb{S}(1)$*, whose distributions are characterized by the following properties:*

— *The process* $(S_i(\overrightarrow{\theta_\infty}))_{i \geq 1}$ *has the same law as the Markov chain* $\tilde{X}$ *started from* 1.

— *Conditionally on* $(S_i(\overrightarrow{\theta_\infty}))_{i \geq 0}$*, the subtrees* $L_i(\overrightarrow{\theta_\infty})$*,* $i \geq 0$ *and* $R_i(\overrightarrow{\theta_\infty})$*,* $i \geq 1$ *are independent random variables, with respective distributions* $\rho_{(S_i(\overrightarrow{\theta_\infty}))}$ *and* $\rho^+_{(S_i(\overrightarrow{\theta_\infty}))}$.

— *We have the joint distributional identities:*

$$\begin{aligned} (S_i(\overleftarrow{\theta_\infty}) - 1)_{i \geq 0} &= (S_i(\overrightarrow{\theta_\infty}))_{i \geq 0} \\ (L_i(\overleftarrow{T_\infty}), \overleftarrow{l_\infty} - 1)_{i \geq 0} &= (R_i(\overrightarrow{\theta_\infty}))_{i \geq 0} \\ (R_i(\overleftarrow{T_\infty}), \overleftarrow{l_\infty} - 1)_{i \geq 0} &= (L_i(\overrightarrow{\theta_\infty}))_{i \geq 0}. \end{aligned}$$

We finally extend this convergence to the associated quadrangulations:

Theorem 1.5. *Let* $\overrightarrow{Q}_\infty = \Phi(\overrightarrow{\theta_\infty})$ *and* $\overleftarrow{Q}_\infty = \Phi(\overleftarrow{\theta_\infty})$*. We have the joint convergence in distribution*

$$(\overrightarrow{Q}^{(k)}_\infty, \overleftarrow{Q}^{(k)}_\infty) \xrightarrow[k \to \infty]{} (\overrightarrow{Q}_\infty, \overleftarrow{Q}_\infty)$$

for the local topology. As a consequence, $Q^{(k)}_\infty$ *converges in distribution towards the quadrangulation of the plane* $\overleftrightarrow{Q}_\infty$ *obtained by gluing together the boundaries of* $\overrightarrow{Q}_\infty$ *and* $\overleftarrow{Q}_\infty$ *in such a way that their root edges are identified.*

Note that Φ is not continuous at points $\overrightarrow{Q}_\infty$ and $\overleftarrow{Q}_\infty$, so this result is not a straightforward consequence of the previous theorem. In the same spirit as Ménard in [52], we have to show that the balls of radius r in $\overrightarrow{Q}_\infty$ and $\overleftarrow{Q}_\infty$ are included into balls of radius $h(r)$ in the corresponding trees with high probability, uniformly in k. This is done in Proposition 4.1.

The distribution of $\overleftrightarrow{Q}_\infty$ could be the subject of further study, in particular concerning its symmetries. Informally, it would be interesting to see if it is invariant under the two following transformations:

— Rerooting $\overleftrightarrow{Q}_\infty$ at the "lowest" edge e belonging to an infinite geodesic $(\gamma(i))_{i\in\mathbb{Z}}$, such that $l(e^-) = 0$ and $l(e^+) = 1$; then taking the quadrangulation obtained by reflection with respect to the root edge.

— Rerooting $\overleftrightarrow{Q}_\infty$ at the "lowest" edge e belonging to an infinite geodesic $(\gamma(i))_{i\in\mathbb{Z}}$, such that $l(e^-) = 0$ and $l(e^+) = 1$; then reversing the root edge.

In the first case, the invariance should be easy to derive from symmetries of the UIPQ. The second question appears more difficult and is work in progress.

The paper is organized in the following way. In Sections 2 and 3, we focus on the convergence of the trees $\theta_\infty^{(k)}$ and $\theta_\infty^{(-k+1)}$. We first give the proof of the convergence of $\theta_\infty^{(k)}$ alone, and then show how the same methods can be applied to derive the joint convergence. Note that the convergence results of Section 2 are not necessary in the proof of the joint convergence, but should make the structure of the proof easier to understand. Finally, Section 4 is devoted to the proof of Theorem 1.5.

2 Convergence of $\theta_\infty^{(k)}$

2.1 Explicit expressions for the distribution of $\theta_\infty^{(k)}$

In this section, we work with a fixed value of $k \in \mathbb{N}$. Let us introduce some notation for particular vertices and subtrees of $\theta_n^{(k)}$, for $n \in \mathbb{N} \cup \{\infty\}$. All the variables we consider also depend on k, and should therefore be denoted with an exponent $^{(k)}$, but we omit it as long as k is fixed, to keep the notation readable. First, let m_n be the graph-distance between $\mathrm{e}_0(\theta_n)$ and $\mathrm{e}_k(\theta_n)$, and $\mathrm{x}_{n,0}, \ldots, \mathrm{x}_{n,m_n}$ denote the sequence of the vertices which appear on the path from $\mathrm{e}_k(\theta_n)$ to $\mathrm{e}_0(\theta_n)$. For all $i \in \{0, \ldots, m_n\}$, let $X_{n,i} = l_n^{(k)}(\mathrm{x}_{n,i})$. We also consider the subtrees which appear on each "side" of the path $(\mathrm{x}_{n,0}, \ldots, \mathrm{x}_{n,m_n})$:

— For all $i \in \{1, \ldots, m_n\}$, let $\tau_{n,i}$ be the subtree of $\theta_n^{(k)}$ containing the vertices v such that in θ_n, we had $\mathrm{x}_{n,i} \leq \mathrm{v} < \mathrm{x}_{n,i-1}$.

— For all $i \in \{0, \ldots, m_n\}$, let $\tau'_{n,i}$ be the subtree of $\theta_n^{(k)}$ containing the vertices v such that in θ_n, we had $\mathrm{v} = \mathrm{x}_{n,i}$, or $\mathrm{x}_{n,i} \prec \mathrm{v}$, $\mathrm{x}_{n,i-1} < \mathrm{v}$ and $\mathrm{x}_{n,i-1} \not\prec \mathrm{v}$.

We emphasize that these subtrees inherit the labels $l_n^{(k)}$ instead of l_n, even if we have to use the orders $<$ and $\prec$ on T_n (instead of $T_n^{(k)}$) to define them. The fact that we have to use these orders may seem a bit clumsy since the subtrees are numbered starting from the root $\mathrm{x}_{n,0}$ of $\theta_n^{(k)}$, but it is necessary to get the distinction between τ_{n,m_n} and τ'_{n,m_n}. Figure 4.4 sums up the above notation.

Our first step is to characterize the joint distribution of m_∞, $(X_{\infty,i})_{0\leq i\leq m_\infty}$, $(\tau_{\infty,i})_{1\leq i\leq m_\infty}$ and $(\tau'_{\infty,i})_{0\leq i\leq m_\infty}$. We introduce some more notation for the sets in which these random variables take their values. For all $m, x, x' \in \mathbb{N}$, let $\mathcal{M}^+_{m,x\to x'}$ denote the set of the walks $(x_1, \ldots, x_m) \in \mathbb{N}^m$ such that $x_1 = x$, $x_m = x'$ and for all $i \leq m-1$, $|x_{i+1} - x_i| \leq 1$. Also let

$$\mathbb{T}_n^+(x) = \{(T, l) \in \mathbb{T}_n(x) : l > 0\} \qquad \forall x \in \mathbb{N},$$
$$\mathbb{T}_n^+ = \bigcup_{x\in\mathbb{N}} \mathbb{T}_n^+(x) \qquad \text{and} \qquad \mathbb{T}^+ = \bigcup_{n\geq 0} \mathbb{T}_n^+.$$

We also use the following facts on the distributions $\rho_{(x)}$ and $\rho^+_{(x)}$: for all $n \geq 0$, it is known that

$$\rho_{(x)}(\{\theta\}) = \frac{1}{2 \cdot 12^n} \qquad \forall x \in \mathbb{Z},\ \theta \in \mathbb{T}_n(x)$$

and Proposition 2.4 of [22] shows that

$$\sum_{n \geq 0} \frac{1}{2 \cdot 12^n} \# \mathbb{T}^+_n(x) = w(x) := \frac{x(x+3)}{(x+1)(x+2)} \qquad \forall x \in \mathbb{N}. \tag{4.4}$$

In particular, for all $n \geq 0$ and $x \in \mathbb{N}$, we have $\rho_{(x)}(\mathbb{T}^+) = w(x)$ and

$$\rho^+_{(x)}(\{\theta\}) = \frac{1}{2w(x)12^n} \qquad \forall \theta \in \mathbb{T}^+_n(x).$$

Finally, for all $m \in \mathbb{N}$ and $(x_0, \ldots, x_m) \in \mathbb{Z}^{m+1}$, we let $\mu_{(x_0,\ldots,x_m)}$ denote the distribution of the forest $(\tilde{\tau}_i)_{0 \leq i \leq m}$ defined as follows. Let I be a uniform random variable in $\{0, \ldots, m\}$. Let $\tilde{\tau}_I$ be a random tree distributed as $(T_\infty, l_\infty + x_I)$, and $\tilde{\tau}_i$, $i \in \{0, \ldots, m\} \setminus \{I\}$ be independent random trees distributed according to $\rho_{(x_i)}$, independent of $\tilde{\tau}_I$.

We can now state the proposition:

Proposition 2.1. *We have $X_{\infty,0} = 0$ a.s., and for all $m \in \mathbb{N}$, $\underline{x} \in \mathcal{M}^+_{m,1\to k}$,*

$$\mathbb{P}(m_\infty = m, (X_{\infty,1}, \ldots, X_{\infty,m}) = \underline{x}) = \frac{m+1}{3^m} \prod_{i=1}^m w(x_i).$$

Moreover, conditionally on $m_\infty = m$ and $(X_{\infty,1}, \ldots, X_{\infty,m}) = \underline{x}$:
— *The forests $(\tau_{\infty,i})_{1 \leq i \leq m}$ and $(\tau'_{\infty,i})_{0 \leq i \leq m}$ are independent.*
— *The trees $\tau_{\infty,i}$, $1 \leq i \leq m$ are independent random variables distributed according to $\rho^+_{(x_i)}$.*
— *The forest $(\tau'_{\infty,i})_{0 \leq i \leq m}$ is distributed according to $\mu_{(0,x_1,\ldots,x_m)}$.*

The proof of this proposition relies on counting the well-labeled trees with n edges such that the corresponding m_n, $(\tau_{n,1}, \ldots, \tau_{n,m})$ take a certain value, and using the convergence (4.2).

Proof. We say that a well-labeled forest with m trees is a m-tuple of well-labeled plane rooted trees $(t_1, \ldots, t_m)$, such that for all $i \in \{1, \ldots, m-1\}$, the labels of the roots of t_i and t_{i+1} differ by at most 1. The number of edges of such a forest is the sum of the numbers of edges of the trees $t_1 \ldots, t_m$. Let $\mathbb{F}_{m,n}$ be the set of well-labeled plane forests with m trees and n edges.

Fix $m, N \geq 0$, $\underline{t} = (t_1, \ldots, t_m) \in \mathbb{F}_{m,N}$ such that the root of t_1 has label 1, and all the labels in $\underline{t}$ are positive. For all $n \in \mathbb{N} \cup \{\infty\}$, let

$$P^{(k)}_n(m, \underline{t}) = \mathbb{P}(m_n = m, (\tau_{n,1}, \ldots, \tau_{n,m}) = \underline{t} | \min l_n \leq -k).$$

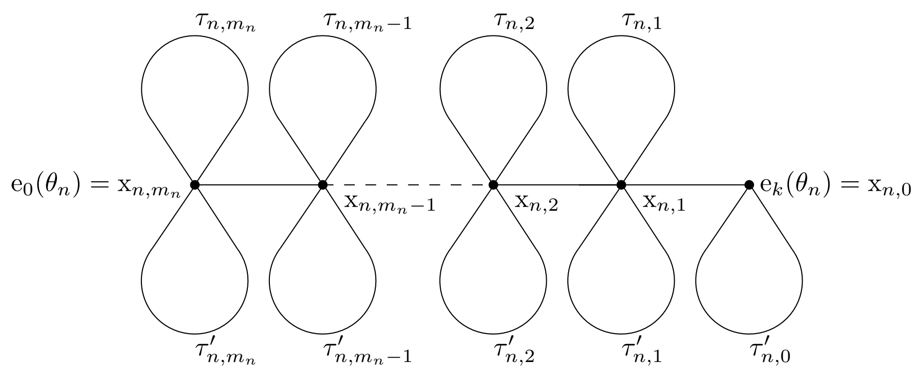

Figure 4.4 – Notation for the vertices and subtrees of $\theta^{(k)}_n$.

We are interested in the behaviour of $P_n^{(k)}(m,\underline{t})$ as $n \to \infty$, for fixed k. Since θ_n is uniform in $\mathbb{T}_n(0)$, we have

$$P_n^{(k)}(m,\underline{t}) = \frac{\mathcal{F}_{m+1,n-(m+N)}}{\#\left\{(T,l) \in \mathbb{T}_n(0) : \min l \leq -k\right\}},$$

where for all $n' \geq 0$,

$$\begin{aligned}\mathcal{F}_{m+1,n'} &= \#\left\{(t'_0,\dots,t'_m) \in \mathbb{F}_{m+1,n'} : \begin{array}{l}\text{the root of } t'_0 \text{ has label 0 and for all } i \geq 1, \\ \text{the root of } t'_0 \text{ has the same label as the root of } t_i\end{array}\right\} \\ &= \#\left\{(t'_0,\dots,t'_m) \in \mathbb{F}_{m+1,n'} : \text{for all } i \geq 1, \text{the root of } t'_0 \text{ has label } 0\right\}.\end{aligned}$$

First note that

$$\#\left(\{(T,l) \in \mathbb{T}_n(0) : \min l \leq -k\}\right) \sim_{n\to\infty} \#\mathbb{T}_n(0).$$

Moreover, it can be seen from the well-known cyclic lemma (see [55]) that

$$\#\mathbb{T}_n(0) = \frac{3^n}{2n+1}\binom{2n+1}{n} \tag{4.5}$$

and

$$\mathcal{F}_{m,n} = \frac{3^n m}{2n+m}\binom{2n+m}{n}. \tag{4.6}$$

Applying these formulas to our case gives

$$\mathcal{F}_{m+1,n-(m+N)} = \frac{3^{n-(m+N)}(m+1)}{2n+1-(m+2N)}\binom{2n+1-(m+2N)}{n-(m+N)},$$

and therefore

$$P_n^{(k)}(m,\underline{t}) \sim_{n\to\infty} \frac{m+1}{3^{m+N}}\binom{2n+1}{n}^{-1}\binom{2n+1-(m+2N)}{n-(m+N)}.$$

We now use Stirling's formula to get an estimate of the binomial coefficients involved:

$$\binom{2n+1}{n} \sim_{n\to\infty} \frac{2\cdot 4^n}{\sqrt{\pi n}},$$

and

$$\binom{2n+1-(m+2N)}{n-(m+N)} \sim_{n\to\infty} \frac{4^n}{2^{m+2N-1}\sqrt{\pi n}}.$$

Putting these together, we obtain

$$P_n^{(k)}(m,\underline{t}) \sim_{n\to\infty} \frac{m+1}{3^{m+N}2^{m+2N}} = \frac{m+1}{6^m 12^N},$$

so the local convergence (4.2) implies that

$$P_\infty^{(k)}(m,\underline{t}) = \frac{m+1}{6^m 12^N}.$$

As a consequence, for all $m \in \mathbb{N}$, $\underline{x} \in \mathcal{M}^+_{m,1\to k}$, we have

$$\mathbb{P}(m_\infty = m, (X_{\infty,1},\dots,X_{\infty,m}) = \underline{x}) = \frac{m+1}{6^m}\prod_{i=1}^m\left(\sum_{n_i\geq 0}\frac{1}{12^{n_i}}\#\mathbb{T}^+_{n_i}(x_i)\right).$$

Recalling equation (4.4), we get

$$\mathbb{P}(m_\infty = m, (X_{\infty,1}, \dots, X_{\infty,m}) = \underline{x}) = \frac{m+1}{6^m} \prod_{i=1}^{m} 2w(x_i) = \frac{m+1}{3^m} \prod_{i=1}^{m} w(x_i).$$

Furthermore, for all $\underline{t} = (t_1, \dots, t_m) \in \mathbb{F}_{m,N}$ such that all the labels in $\underline{t}$ are positive, the conditional probability

$$\mathbb{P}((\tau_{\infty,1}, \dots, \tau_{\infty,m}) = \underline{t} \mid m_\infty = m, (X_{\infty,1}, \dots, X_{\infty,m}) = \underline{x})$$

is equal to

$$\frac{m+1}{6^m 12^N} \cdot \left(\frac{m+1}{3^m} \prod_{i=1}^{m} w(x_i) \right)^{-1} = \prod_{i=1}^{m} \frac{1}{2w(x_i)12^{|t_i|}} = \prod_{i=1}^{m} \rho^{+}_{(x_i)}(t_i),$$

hence the conditional distribution of $(\tau_{\infty,1}, \dots, \tau_{\infty,m_\infty})$.

Finally, conditionally on $m_\infty = m$, $(X_{\infty,1}, \dots, X_{\infty,m}) = \underline{x}$ and $(\tau_{\infty,1}, \dots, \tau_{\infty,m}) = \underline{t} \in \mathbb{F}_{m,N}$, the trees $(\tau'_{n,0}, \dots, \tau'_{n,m})$ form a uniform labeled forest with $m+1$ trees and $n-m-N$ edges, hence the distribution of the limit given in the statement. □

To get the limit of $\theta_\infty^{(k)}$, the main step will consist in showing that for any $r \in \mathbb{N}$, the labels $X_{\infty,1}^{(k)}, \dots, X_{\infty,r}^{(k)}$ converge in distribution to the r first steps of the Markov chain $\tilde{X}$ started at 1, as $k \to \infty$. For the moment, we show how to make $\tilde{X}$ appear in the above expression; the fact that it is indeed the limit is the purpose of Proposition 2.4.

We first introduce the random walk $(\hat{X}_i)_{i\geq 0}$ with uniform random steps in $\{-1, 0, 1\}$. From now on, we also adopt the usual notation $\mathbb{E}_x[\,\cdot\,]$ for the conditional expectation $\mathbb{E}[\,\cdot\,|\hat{X}_0 = x]$, for all x. The expression of the lemma implies that

$$\mathbb{P}(m_\infty = m) = \frac{m+1}{3} \mathbb{E}_1 \left[\prod_{i=0}^{m-1} w(\hat{X}_i) \mathbf{1}_{\{\hat{X}_{m-1} = k\}} \right].$$

(Note that the term in the expectation is zero if we do not have $\hat{X}_i \geq 1$ for all $i \leq m-1$.) Let $f(x) = x(x+3)(2x+3)$ for all $x \in \mathbb{R}$, and

$$M_j = \frac{f(\hat{X}_j)}{f(\hat{X}_0)} \prod_{i=0}^{j-1} w(\hat{X}_i) \qquad \forall j \geq 0.$$

Under the assumption $\hat{X}_0 = 1$, the process $(M_i)_{i\geq 0}$ is a martingale. Using this new process, we get

$$\begin{aligned} \mathbb{P}(m_\infty = m) &= \frac{m+1}{3} \mathbb{E}_1 \left[\frac{f(1)w(k)}{f(k)} M_{m-1} \mathbf{1}_{\{\hat{X}_{m-1} = k\}} \right] \\ &= \frac{f(1)w(k)}{3f(k)} (m+1) \mathbb{P}_1(\tilde{X}_{m-1} = k), \end{aligned}$$

where $\tilde{X}$ is defined as the image of $\hat{X}$ under the measure-change given by the martingale M, i.e. the Markov process such that $\mathbb{E}[\phi(\tilde{X}_i)] = \mathbb{E}[M_i \phi(\hat{X}_i)]$ for every continuous bounded function ϕ. Computing the transition probabilities of $\tilde{X}$ gives:

$$\begin{aligned} p_x = \mathbb{P}_x(\tilde{X}_1 = x+1) &= \frac{f(x+1)w(x)}{3f(x)} \\ r_x = \mathbb{P}_x(\tilde{X}_1 = x) &= \frac{w(x)}{3} \\ q_x = \mathbb{P}_x(\tilde{X}_1 = x-1) &= \frac{f(x-1)w(x)}{3f(x)}, \end{aligned}$$

hence the expressions given in the Introduction.

2.2 Two useful quantities

To prove of the convergence of $\theta_\infty^{(k)}$, we will need estimates for the quantities

$$H_x(k) = \sum_{m\geq 1} \mathbb{P}_x(\tilde{X}_{m-1} = k)$$
$$H_x^*(k) = \sum_{m\geq 1} (m+1)\mathbb{P}_x(\tilde{X}_{m-1} = k),$$

depending on the values of $k, x \in \mathbb{N}$. In practice, these estimates are best obtained through explicit computation; the expressions we get are given in the two following lemmas. We use the notations $\tilde{T}_y = \inf\{t \geq 1 : \tilde{X}_t = y\}$ and $h(y) = y(y+1)(y+2)(y+3)(2y+3)$, for all $y \in \mathbb{N}$. (In this section, we mainly work on Markov processes, and use the letters t and T for associated times instead of trees.)

Lemma 2.2. *Fix $k \geq 2$, $x \in \mathbb{N}$. We have the following equalities:*

— *if $x \leq k$,*

$$H_x(k) = \frac{3}{10}(2k+3),$$

— *if $x > k$,*

$$H_x(k) = \frac{3h(k)}{10h(x)}(2k+3).$$

Proof. Fix $x, k \in \mathbb{N}$. First note that we can write $H_x(k)$ as

$$H_x(k) = \sum_{m\geq 0} \mathbb{P}_x(\tilde{X}_m = k) = \mathbb{E}_x\left[\sum_{m\geq 0} \mathbf{1}_{\{\tilde{X}_m = k\}}\right] = \mathbf{1}_{\{x=k\}} + \mathbb{E}_x\left[\mathbf{1}_{\{\tilde{T}_k<\infty\}} \sum_{m\geq 0} \mathbf{1}_{\{\tilde{X}_{\tilde{T}_k+m} = k\}}\right].$$

Now, applying the Markov property at the stopping time $\tilde{T}_k$ yields

$$H_x(k) = \mathbf{1}_{\{x=k\}} + \mathbb{P}_x(\tilde{T}_k < \infty)\mathbb{E}_k\left[\sum_{m\geq 0} \mathbf{1}_{\{\tilde{X}_m = k\}}\right].$$

For all $y \geq 0$, let

$$K_{y+1,j} = \frac{q_{y+1}\cdots q_{y+j}}{p_{y+1}\cdots p_{y+j}} = \prod_{z=y+1}^{y+j} \frac{f(z-1)}{f(z+1)} = \frac{f(y)f(y+1)}{f(y+j)f(y+j+1)}.$$

Since $K_{1,j}$ is the general term of a converging series, the Markov chain $\tilde{X}$ is transient, and as a consequence, we have

$$H_x(k) = \mathbf{1}_{\{x=k\}} + \frac{\mathbb{P}_x(\tilde{T}_k < \infty)}{\mathbb{P}_k(\tilde{T}_k = \infty)}. \tag{4.7}$$

To compute these quantities, it is enough to know the expression of $\mathbb{P}_{y+1}(\tilde{T}_y = \infty)$ for all $y \geq k$, which is a well-known property of birth-and-death processes:

$$\mathbb{P}_{y+1}(\tilde{T}_y = \infty) = \frac{1}{\sum_{j\geq 0} K_{y+1,j}}$$

Computing the sum $\sum_{j\geq 0} K_{y+1,j}$ yields

$$\mathbb{P}_{y+1}(\tilde{T}_y = \infty) = \frac{10(y+2)}{(y+4)(2y+5)}. \tag{4.8}$$

As a consequence, we get the following results:

— If $x < k$, then

$$\mathbb{P}_x(\tilde{T}_k < \infty) = 1.$$

— If $x = k$, then

$$\mathbb{P}_x(\tilde{T}_k < \infty) = 1 - p_k \mathbb{P}_{k+1}(\tilde{T}_k = \infty) = \frac{6k-1}{3(2k+3)}.$$

— If $x > k$, then

$$\mathbb{P}_x(\tilde{T}_k < \infty) = \prod_{y=k}^{x-1} \mathbb{P}_{y+1}(\tilde{T}_y < \infty) = \frac{h(k)}{h(x)}.$$

Together with (4.7), this completes the proof of the lemma. $\square$

Note that the values we obtain can also be computed using the recurrence relations

$$\begin{cases} H_1(1) = 1 + r_1 H_1(1) + q_2 H_1(2) \\ H_1(k) = p_{k-1} H_1(k-1) + r_k H_1(k) + q_{k+1} H_1(k+1) \qquad \forall k \geq 2 \end{cases} \tag{4.9}$$

and, for all $k \in \mathbb{N}$,

$$\begin{cases} H_1(k) = \mathbf{1}_{\{k=1\}} + r_1 H_1(k) + p_1 H_2(k) \\ H_x(k) = \mathbf{1}_{\{k=x\}} + p_x H_{x+1}(k) + r_x H_x(k) + q_x H_{x-1}(k) \qquad \forall x \geq 2, \end{cases} \tag{4.10}$$

which stem from the Markov property of $\tilde{X}$. Nevertheless, we would still have to go through part of the previous calculations to get the value of $H_1(1)$. In the proof of the following lemma, we will find it easier to use this approach.

Lemma 2.3. *Fix $k \geq 2$, $x \in \mathbb{N}$, and let $C_x = \frac{3}{14}((x+1)(x+2) - 6)$. We have the following equalities:*

— *if $x < k$,*

$$H_x^*(k) = \frac{3f(k)}{f(1)w(k)} - \frac{3C_x}{10}(2k+3),$$

— *if $x \geq k$,*

$$H_x^*(k) = \frac{3f(k)}{f(1)w(k)} - \frac{3}{10}(2k+3)\left(C_x + 1 - \frac{h(k)}{h(x)}\right).$$

Proof. The first step of the proof consists in computing $H_1^*(1)$. We will then obtain $H_x^*(k)$ as the unique solution of recursive systems having this initial value. Note that since $\mathbb{P}_1(\tilde{X}_0 = 1) = 1$, we have

$$H_1^*(1) = 2 + \sum_{m \geq 1} (m+2) \mathbb{P}_1(\tilde{X}_m = 1).$$

Let us rewrite the second term using the first return time in 1, as in the proof of the previous lemma:

$$\begin{aligned} H_1^*(1) &= 2 + \sum_{t \geq 1} \mathbb{P}_1(\tilde{T}_1 = t) \sum_{m \geq t} (m+2) \mathbb{P}_1(\tilde{X}_m = 1 \mid \tilde{T}_1 = t) \\ &= 2 + \sum_{t \geq 1} \mathbb{P}_1(\tilde{T}_1 = t) \sum_{m \geq 0} (m+t+2) \mathbb{P}_1(\tilde{X}_m = 1) \\ &= 2 + H_1^*(1) \sum_{t \geq 1} \mathbb{P}_1(\tilde{T}_1 = t) + H_1(1) \sum_{t \geq 1} t \mathbb{P}_1(\tilde{T}_1 = t) \\ &= 2 + H_1^*(1) \mathbb{P}_1(\tilde{T}_1 < \infty) + H_1(1) \mathbb{E}_1\left[\tilde{T}_1 \mathbf{1}_{\{\tilde{T}_1 < \infty\}}\right]. \end{aligned}$$

Thus, we have

$$H_1^*(1) = \frac{1}{\mathbb{P}_1(\tilde{T}_1 = \infty)} \left(2 + \frac{3}{2} \mathbb{E}_1[\tilde{T}_1 \mid \tilde{T}_1 < \infty] \mathbb{P}_1(\tilde{T}_1 < \infty) \right).$$

Using the value of $\mathbb{P}_1(\tilde{T}_1 < \infty)$ obtained in the previous proof, we get

$$H_1^*(1) = \frac{3}{2} \left(2 + \frac{1}{2} \mathbb{E}_1[\tilde{T}_1 \mid \tilde{T}_1 < \infty] \right). \tag{4.11}$$

To work out the value of the above expectation, we study the process $\tilde{X}^*$ having the law of $\tilde{X}$ conditioned on returning to 1 infinitely often. This process is a recurrent Markov chain whose transition probabilities can be computed explicitly. Indeed, letting

$$\begin{aligned} p_x^* &:= \mathbb{P}_x(\tilde{X}_1 = x+1 \mid \tilde{T}_1 < \infty) \\ r_x^* &:= \mathbb{P}_x(\tilde{X}_1 = x \mid \tilde{T}_1 < \infty) \\ q_x^* &:= \mathbb{P}_x(\tilde{X}_1 = x-1 \mid \tilde{T}_1 < \infty), \end{aligned}$$

Bayes' law yields

$$\begin{aligned} p_x^* &= \frac{\mathbb{P}_x(\tilde{X}_1 = x+1)\mathbb{P}_x(\tilde{T}_1 < \infty \mid \tilde{X}_1 = x+1)}{\mathbb{P}_x(\tilde{T}_1 < \infty)} = \frac{p_x \mathbb{P}_{x+1}(\tilde{T}_1 < \infty)}{\mathbb{P}_x(\tilde{T}_1 < \infty)}, \\ r_x^* &= \frac{\mathbb{P}_x(\tilde{X}_1 = x)\mathbb{P}_x(\tilde{T}_1 < \infty \mid \tilde{X}_1 = x)}{\mathbb{P}_x(\tilde{T}_1 < \infty)} = \begin{cases} r_x \text{ if } x \neq 1 \\ \frac{r_1}{\mathbb{P}_1(\tilde{T}_1 < \infty)} \text{ if } x = 1, \end{cases} \\ q_x^* &= \frac{\mathbb{P}_x(\tilde{X}_1 = x-1)\mathbb{P}_x(\tilde{T}_1 < \infty \mid \tilde{X}_1 = x-1)}{\mathbb{P}_x(\tilde{T}_1 < \infty)} = \begin{cases} \frac{q_x \mathbb{P}_{x-1}(\tilde{T}_1 < \infty)}{\mathbb{P}_x(\tilde{T}_1 < \infty)} \text{ if } x \neq 2 \\ \frac{q_2}{\mathbb{P}_2(\tilde{T}_1 < \infty)} \text{ if } x = 1. \end{cases} \end{aligned}$$

Note that, for all $x \geq 2$,

$$\mathbb{P}_{x+1}(\tilde{T}_1 < \infty) = \mathbb{P}_{x+1}(\tilde{T}_x < \infty)\mathbb{P}_x(\tilde{T}_1 < \infty),$$

so we can again use equation (4.8). Finally, we get $p_1^* = \frac{1}{3}$, $r_1^* = \frac{2}{3}$, $q_1^* = 0$, and for all $x \geq 2$

$$\begin{aligned} p_x^* &= \frac{x}{3(x+2)} \\ r_x^* &= \frac{x(x+3)}{3(x+1)(x+2)} \\ q_x^* &= \frac{x+3}{3(x+1)}. \end{aligned}$$

To get the value of $\mathbb{E}_1[\tilde{T}_1 \mid \tilde{T}_1 < \infty]$, it is now enough to compute the invariant measure Π of $\tilde{X}^*$. We do so by using reversibility: the detailed balanced equation $\Pi(x)p_x^* = \Pi(x+1)q_{x+1}^*$ implies

$$\frac{\Pi(x+1)}{\Pi(x)} = \begin{cases} \frac{x}{x+4} \text{ if } x \geq 2 \\ \frac{3}{5} \text{ if } x = 1. \end{cases}$$

As a consequence,

$$\sum_{x \geq 1} \Pi(x) = \Pi(1) \left(1 + \frac{3}{5} \sum_{x \geq 2} \frac{2 \times 3 \times 4 \times 5}{x(x+1)(x+2)(x+3)} \right) = \Pi(1) \left(1 + \frac{3}{5} \times 120 \times \frac{1}{72} \right),$$

so Π is a probability measure if and only if $\Pi(1) = \frac{1}{2}$. This implies

$$\mathbb{E}_1[\tilde{T}_1 \mid \tilde{T}_1 < \infty] = \frac{1}{\Pi(1)} = 2.$$

Injecting this value into (4.11) gives $H_1^*(1) = \frac{9}{2}$.

For the second step, we keep $x = 1$, and compute the values of $H_1^*(k)$ for $k \in \mathbb{N}$. As above, we first shift indices and set the first term aside:

$$H_1^*(k) = 2\mathbf{1}_{\{k=1\}} + \sum_{m\geq 0}(m+3)\mathbb{P}_1(\tilde{X}_{m+1} = k).$$

Applying the Markov property at time m in each of the terms gives the following recurrence relations:

— For $k = 1$,

$$\begin{aligned} H_1^*(1) &= 2 + \sum_{m\geq 0}(m+3)\left(r_1\mathbb{P}_1(\tilde{X}_m = 1) + q_2\mathbb{P}_1(\tilde{X}_m = 2)\right) \\ &= 2 + r_1 H_1^*(1) + q_2 H_1^*(2) + H_1(1) - 1. \end{aligned}$$

— For all $k \geq 2$,

$$\begin{aligned} H_1^*(k) &= \sum_{m\geq 0}(m+3)\left(p_{k-1}\mathbb{P}_1(\tilde{X}_m = k-1) + r_k\mathbb{P}_1(\tilde{X}_m = k) + q_{k+1}\mathbb{P}_1(\tilde{X}_m = k+1)\right) \\ &= p_{k-1}(H_1^*(k-1) + H_1(k-1)) + r_k(H_1^*(k) + H_1(k)) + q_{k+1}(H_1^*(k+1) + H_1(k+1)) \\ &= p_{k-1}H_1^*(k-1) + r_k H_1^*(k) + q_{k+1}H_1^*(k+1) + H_1(k). \end{aligned}$$

(Note that we have used implicitly the fact that $H_1(k)$ verifies the similar system (4.9)). Using the values obtained in Lemma 2.2, we get the recursive system

$$\begin{cases} H_1^*(1) = \frac{9}{2} \\ H_1^*(2) = \frac{1}{q_2}\left((1-r_1)H_1^*(1) - \frac{5}{2}\right) \\ H_1^*(k+1) = \frac{1}{q_{k+1}}\left((1-r_k)H_1^*(k) - p_{k-1}H_1^*(k-1) - \frac{3}{10}(2k+3)\right). \end{cases}$$

It is now easy to check that $\frac{3f(k)}{f(1)w(k)}$ is also a solution of this system, and therefore is equal to $H_1^*(k)$.

In the third and last step, we fix the value of k, and write recurrence relations for $H_x^*(k)$, $x \in \mathbb{N}$. To this end, we again use the Markov property, but at time 1 (with the convention that $H_0^*(k) = 0$, to keep the setting general):

$$\begin{aligned} H_x^*(k) &= 2\mathbf{1}_{\{x=k\}} + \sum_{m\geq 0}(m+3)\mathbb{P}_x(\tilde{X}_{m+1} = k) \\ &= 2\mathbf{1}_{\{x=k\}} + \sum_{m\geq 0}(m+3)(p_x\mathbb{P}_{x+1}(\tilde{X}_m = k) + r_x\mathbb{P}_x(\tilde{X}_{m+1} = k) + q_x\mathbb{P}_{x-1}(\tilde{X}_m = k)) \\ &= 2\mathbf{1}_{\{x=k\}} + p_x H_{x+1}^*(k) + r_x H_x^*(k) + q_x H_{x-1}^*(k) + p_x H_{x+1}(k) + r_x H_x(k) + q_x H_{x-1}(k) \\ &= \mathbf{1}_{\{x=k\}} + p_x H_{x+1}^*(k) + r_x H_x^*(k) + q_x H_{x-1}^*(k) + H_x(k). \end{aligned}$$

This gives the system

$$\begin{cases} H_1^*(k) = \frac{3f(k)}{f(1)w(k)} \\ H_{x+1}^*(k) = \frac{1}{p_x}\left((1-r_x)H_x^*(k) - q_x H_{x-1}^*(k) - H_x(k) - \mathbf{1}_{\{x=k\}}\right). \end{cases}$$

We first solve these equations for $x < k$, so that the last term is zero. The solution is of the form given in the lemma if and only if C_x is such that

$$\begin{cases} C_0 = C_1 = 0 \\ C_{x+1} = \frac{1}{p_x}((1-r_x)C_x - q_x C_{x-1} + 1). \end{cases}$$

This is indeed the case for $C_x = \frac{3}{14}((x+1)(x+2) - 6)$. Now, for $x \geq k$, we seek a solution of the form

$$H_x^*(k) = \frac{3f(k)}{f(1)w(k)} - \frac{3C_x}{10}(2k+3) - C'_{k,x}.$$

The recursive system can be translated into $C'_{k,k} = 0$, $C'_{k,k+1} = \frac{1}{p_x}$ and

$$C'_{k,x+1} = \frac{1}{p_x}((1-r_x)C'_{k,x} - q_x C'_{k,x-1}),$$

or equivalently

$$p_x(C'_{k,x+1} - C'_{k,x}) = q_x(C'_{k,x} - C'_{k,x-1}).$$

Thus, for $x \geq k+1$, we get

$$\begin{aligned} C'_{k,x} &= \sum_{y=k}^{x-1} \frac{q_{k+1}\dots q_y}{p_{k+1}\dots p_y}\frac{1}{p_k} \\ &= \sum_{y=k}^{x-1} \frac{f(k)f(k+1)}{f(y)f(y+1)}\frac{1}{p_k} \\ &= \frac{f(k)f(k+1)}{p_k}\frac{h(x)-h(k)}{10h(k)h(x)}. \end{aligned}$$

Using the expressions of f, h and p_k, we conclude that

$$\begin{aligned} &= \frac{(k+4)(2k+5)}{10(k+2)p_k}\left(1 - \frac{h(k)}{h(x)}\right) \\ &= \frac{3(2k+3)}{10}\left(1 - \frac{h(k)}{h(x)}\right). \end{aligned}$$

This ends the proof. □

2.3 Proof of the convergence

We are now ready to give the proof of the convergence of $\theta_\infty^{(k)}$. We begin with the convergence of the labels $(X_{\infty,i}^{(k)})$ towards the Markov chain $\tilde{X}$.

Proposition 2.4. *Fix $r \in \mathbb{N}$. For any continuous bounded function F from $\mathbb{R}^r$ into $\mathbb{R}$, we have*

$$\mathbb{E}[F(X_{\infty,1}^{(k)}, \dots, X_{\infty,r}^{(k)})] \xrightarrow[k\to\infty]{} \mathbb{E}_1[F(\tilde{X}_0, \dots, \tilde{X}_{r-1})].$$

Proof. Let $k \geq r$. The computations of Section 2.1 show that

$$\mathbb{E}[F(X_{\infty,1}^{(k)}, \dots, X_{\infty,r}^{(k)})] = \sum_{m\geq 1} \frac{m+1}{3}\mathbb{E}_1\left[F(\hat{X}_0, \dots, \hat{X}_{r-1})\mathbf{1}_{\{\hat{X}_{m-1}=k\}} M_{m-1}\frac{f(1)w(k)}{f(k)}\right].$$

Since $k \geq r$, the term $\mathbf{1}_{\{\hat{X}_{m-1}=k\}}$ is zero for $m < r$. Applying the Markov property allows us to write $\mathbb{E}[F(X^{(k)}_{\infty,1}, \ldots, X^{(k)}_{\infty,r})]$ as

$$\sum_{m \geq r} \frac{m+1}{3} \mathbb{E}_1 \left[F(\hat{X}_0, \ldots, \hat{X}_{r-1}) \frac{f(1)}{f(\hat{X}_{r-1})} M_{r-1} \mathbb{E}_{\hat{X}_{r-1}} \left[\mathbf{1}_{\{\hat{X}'_{m-r}=k\}} M'_{m-r} \frac{f(\hat{X}_{r-1}) w(k)}{f(k)} \right] \right],$$

where $\hat{X}'$ is an independent copy of the process $\hat{X}$, and for all $j \in \mathbb{N}$

$$M'_j = \frac{f(\hat{X}'_j)}{f(\hat{X}'_0)} \prod_{i=0}^{j-1} w(\hat{X}'_i).$$

Therefore, $\mathbb{E}[F(X^{(k)}_{\infty,1}, \ldots, X^{(k)}_{\infty,r})]$ is equal to

$$\mathbb{E}_1 \left[M_{r-1} F(\hat{X}_0, \ldots, \hat{X}_{r-1}) \frac{f(1) w(k)}{3 f(k)} \sum_{m \geq 0} (m + r + 1) \mathbb{E}_{\hat{X}_{r-1}} \left[\mathbf{1}_{\{\hat{X}'_m = k\}} M'_m \right] \right].$$

Since $\hat{X}_{r-1} \leq r$ a.s., we now have to estimate the sum $\sum_{m \geq 0} (m + r + 1) \mathbb{E}_x \left[\mathbf{1}_{\{\hat{X}'_m = k\}} M'_m \right]$, for all $x \leq r$. We first express this quantity using $H^*_x(k)$ and $H_x(k)$:

$$\begin{aligned} \sum_{m \geq 0} (m + r + 1) \mathbb{E}_x \left[\mathbf{1}_{\{\hat{X}'_m = k\}} M'_m \right] &= \sum_{m \geq 0} (m + r + 1) \mathbb{P}_x (\tilde{X}'_m = k) \\ &= H^*_x(k) + (r - 1) H_x(k). \end{aligned}$$

Now, the results of Lemmas 2.2 and 2.3 yield

$$\sum_{m \geq 0} (m + r + 1) \mathbb{E}_x \left[\mathbf{1}_{\{\hat{X}'_m = k\}} M'_m \right] = \frac{3 f(k)}{f(1) w(k)} + (r - 1 - C_x) \frac{3}{10} (2k + 3).$$

As a consequence, we have

$$\frac{f(1) w(k)}{3 f(k)} \sum_{m \geq 0} (m + r + 1) \mathbb{E}_x \left[\mathbf{1}_{\{\hat{X}'_m = k\}} M'_m \right] \xrightarrow[k \to \infty]{} 1$$

uniformly in $x \leq r$, hence the result. □

Note that we have only used part of the results of Lemmas 2.2 and 2.3 (namely, the case where $x \leq k$). The remaining expressions will play a role in the proof of the joint convergence.

The convergence of $\theta^{(k)}_\infty$ towards $\overrightarrow{\theta_\infty}$ can now be obtained by putting together the results of Proposition 2.1 and Proposition 2.4. Indeed, letting $I^{(k)}$ denote the unique index i such that $\tau'_{\infty,i}$ is infinite, conditionally on $I^{(k)} \geq r$, we have that:

— The points $\mathfrak{s}_i(\theta^{(k)}_\infty)$ and $\mathrm{x}^{(k)}_{\infty,i}$ are the same for all $i \leq r$, hence the equalities $R_i(\theta^{(k)}_\infty) = \tau^{(k)}_{\infty,i}$ and $L_i(\theta^{(k)}_\infty) = (\tau'_{\infty,i})^{(k)}$ for all $i < r$.

— As a consequence, $(S_i(\theta^{(k)}_\infty))_{1 \leq i \leq r}$ converges in distribution to $(\tilde{X}_i)_{0 \leq i \leq r-1}$ for $\tilde{X}_0 = 1$.

— Conditionally on $(S_i(\theta^{(k)}_\infty))_{0 \leq i < r}$, the subtrees $L_i(\theta^{(k)}_\infty)$, $0 \leq i < r$ and $R_i(\theta^{(k)}_\infty)$, $1 \leq i < r$ are independent random variables, with respective distributions $\rho_{(S_i(\theta^{(k)}_\infty))}$ and $\rho^+_{(S_i(\theta^{(k)}_\infty))}$.

Since

$$\mathbb{P}(I^{(k)} < r) = \mathbb{E} \left[\frac{r}{m^{(k)}_\infty + 1} \right] \leq \frac{r}{k+1} \xrightarrow[k \to \infty]{} 0,$$

this gives the desired convergence.

3 Joint convergence of $(\theta_\infty^{(k)}, \theta_\infty^{(-k+1)})$

3.1 Explicit expressions for the joint distribution

As in the previous section, we first fix k, and use the convergence of $(\theta_n^{(k)}, \theta_n^{(-k+1)})$ to study $(\theta_\infty^{(k)}, \theta_\infty^{(-k+1)})$. Let $n \in \mathbb{N} \cup \{\infty\}$. We introduce some new notation, summed-up in Figure 4.5. To simplify what follows, we write e_0, e_k, e_{-k+1} and v_k instead of $\mathrm{e}_0(\theta_n)$, $\mathrm{e}_k(\theta_n)$, $\mathrm{e}_{-k+1}(\theta_n)$ and $\mathrm{v}_k(\theta_n)$.

We first deal with the branches between e_0, e_k and e_{-k+1}. Let $a_n = d_n(\mathrm{v}_k, \mathrm{e}_k)$, $b_n = d_n(\mathrm{v}_k, \mathrm{e}_{-k+1})$ and $c_n = d_n(\mathrm{e}_0, \mathrm{v}_k)$, where d_n denotes the graph-distance on θ_n. Let $\mathrm{x}_{n,0}, \ldots, \mathrm{x}_{n,a_n}$ be the vertices on the path from e_k to v_k, $\mathrm{y}_{n,0}, \ldots, \mathrm{y}_{n,b_n}$ the ones on the path from e_{-k+1} to v_k, and $\mathrm{z}_{n,0}, \ldots, \mathrm{z}_{n,c_n}$ the ones on the path from v_k to e_0. For the corresponding labels, we use capital letters: $X_{n,i} = l_n^{(k)}(\mathrm{x}_{n,i})$, $Y_{n,i} = l_n^{(k)}(\mathrm{y}_{n,i})$ and $Z_{n,i} = l_n^{(k)}(\mathrm{z}_{n,i})$ for all i.

We now add notation for the subtrees which are grafted on these branches. Again, we use the orders $\prec$ and $<$ on the vertices of θ_n in these definitions, even if we think of these trees as subtrees of $\theta_n^{(k)}$ (in particular, they inherit the labels $l_n^{(k)}$).

— For all $i \in \{1, \ldots, a_n + c_n\}$, let $\tau_{n,i}$ be the subtree containing the vertices v such that :
  — if $i \leq a_n$, then $\mathrm{x}_{n,i} \leq \mathrm{v} < \mathrm{x}_{n,i-1}$,
  — if $a_n + 1 \leq i \leq a_n + c_n$, then $\mathrm{z}_{n,i-a_n} \leq \mathrm{v} < \mathrm{z}_{n,i-a_n-1}$.
— For all $i \in \{1, \ldots, b_n + c_n\}$, let $\tau'_{n,i}$ be the subtree containing the vertices v such that :
  — if $i \leq b_n$, either $\mathrm{v} = \mathrm{y}_{n,i}$, or $\mathrm{y}_{n,i} \prec \mathrm{v}$, $\mathrm{y}_{n,i-1} < \mathrm{v}$ and $\mathrm{y}_{n,i-1} \not\prec \mathrm{v}$,
  — if $b_n + 1 \leq i \leq b_n + c_n$, either $\mathrm{v} = \mathrm{z}_{n,i-b_n}$, or $\mathrm{z}_{n,i-b_n} \prec \mathrm{v}$, $\mathrm{z}_{n,i-b_n-1} < \mathrm{v}$ and $\mathrm{z}_{n,i-b_n-1} \not\prec \mathrm{v}$.
— For all $i \in \{0, \ldots, a_n\}$, let $\overline{\tau}_{n,i}$ be the subtree containing the vertices v such that :
  — if $i = 0$, then $\mathrm{x}_0 \preceq \mathrm{v}$,
  — otherwise, either $\mathrm{v} = \mathrm{x}_{n,i}$, or $\mathrm{x}_{n,i} \prec \mathrm{v}$, $\mathrm{x}_{n,i-1} < \mathrm{v}$ and $\mathrm{x}_{n,i-1} \not\prec \mathrm{v}$.
— For all $i \in \{0, \ldots, b_n - 1\}$, let $\overline{\tau}'_{n,i}$ be the subtree containing the vertices v such that :
  — if $i = 0$, then $\mathrm{y}_0 \preceq \mathrm{v}$.
  — otherwise, $\mathrm{y}_{n,i} \leq \mathrm{v} < \mathrm{y}_{n,i-1}$,

As in section 2.1, for all these variables, there should be an exponent $^{(k)}$ in the notation, but we omit this precision as long as k remains constant.

Fix $a, b, c, N, N' \geq 0$, $\underline{t} = (t_1, \ldots, t_{a+c}) \in \mathbb{F}_{a+c,N}$ and $\underline{t'} = (t'_1, \ldots, t'_{b+c}) \in \mathbb{F}_{b+c,N'}$ such that:
— the root of t_1 has label 1, and all labels in $\underline{t}$ are positive,
— the root of t'_1 has label 2, and all labels in $\underline{t'}$ are greater than 1,
— for all $i \leq c$, the labels of the roots of t_{a+c-i} and t'_{b+c-i} are the same.

Let

$$P_n^{(k)}(a, b, c, \underline{t}, \underline{t'}) = \mathbb{P}(a_n = a, b_n = b, c_n = c, (\tau_1, \ldots, \tau_{a+c}) = \underline{t}, (\tau'_1, \ldots, \tau'_{b+c}) = \underline{t'}).$$

We are once again interested in the behaviour of $P_n^{(k)}(a, b, c, \underline{t}, \underline{t'})$ as $n \to \infty$, for fixed k.

Lemma 3.1. *Using the above notation, we have*

$$P_n^{(k)}(a, b, c, \underline{t}, \underline{t'}) \sim_{n \to \infty} \frac{a+b+1}{6^{a+b} 12^{c+N+N'}}.$$

Proof. Recall that $\mathcal{F}_{m,n'}$ denotes the number of well-labeled forests with m trees, n' edges and prescribed root labels. Since θ_n in uniform in $\mathbb{T}_n(0)$, we have

$$P_n^{(k)}(a, b, c, \underline{t}, \underline{t'}) = \frac{\mathcal{F}_{a+b+1, n-(a+b+c+N+N')}}{\#\left\{(T, l) \in \mathbb{T}_n(0) : \min_{V(T)} l \leq -k\right\}}.$$

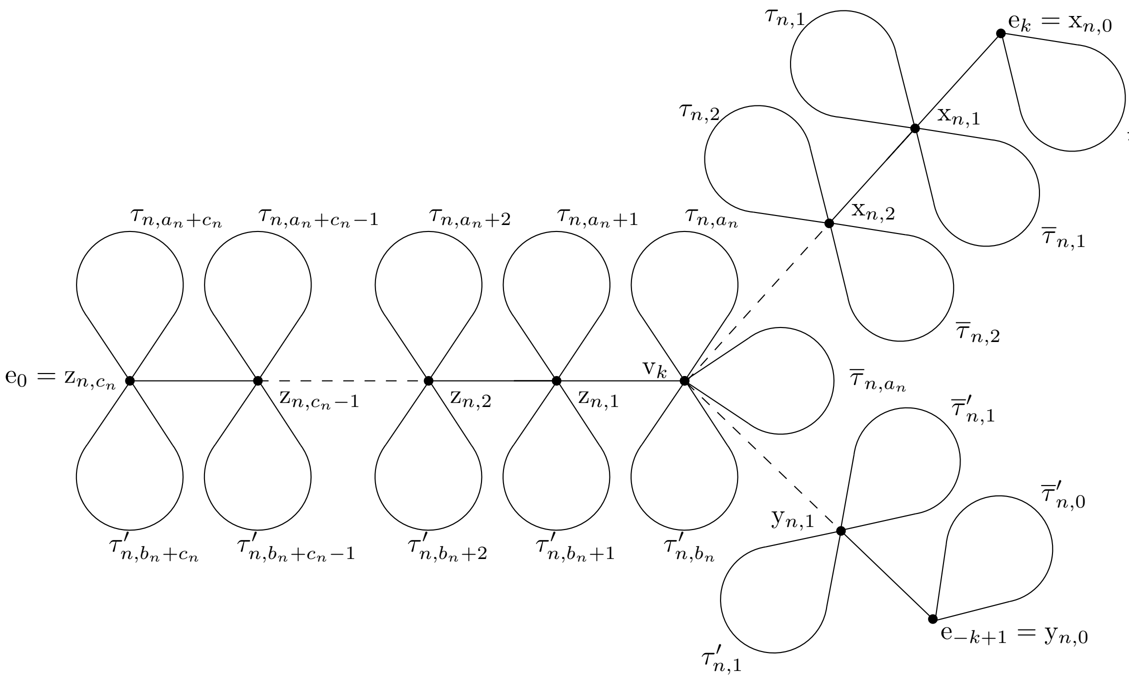

Figure 4.5 – Notation for the vertices and subtrees of θ_n, with distinguished points e_k and e_{-k+1}.

Using equations (4.5) and (4.6) yields

$$\mathcal{F}_{a+b+1,n-(a+b+c+N+N')} = \frac{3^{n-(a+b+c+N+N')}(a+b+1)}{2n+1-(a+b+2c+2N+2N')} \binom{2n+1-(a+b+2c+2N+2N')}{n-(a+b+c+N+N')},$$

and

$$P_n^{(k)}(a,b,c,\underline{t},\underline{t'}) \sim_{n\to\infty} \frac{a+b+1}{3^{a+b+c+N+N'}} \binom{2n+1}{n}^{-1} \binom{2n+1-(a+b+2c+2N+2N')}{n-(a+b+c+N+N')}.$$

Stirling's formula now gives

$$P_n^{(k)}(a,b,c,\underline{t},\underline{t'}) \sim_{n\to\infty} \frac{a+b+1}{3^{a+b+c+N+N'}2^{a+b+2c+2N+2N'}},$$

hence the lemma. □

Recall that for all $m, x, x' \in \mathbb{N}$, $\mathbb{T}_m^+(x)$ is the set of the labeled trees $(T,l) \in \mathbb{T}_m$ such that $l > 0$ and the root of T has label x, and $\mathcal{M}_{m,x\to x'}^+$ is the set of the walks $(x_1,\dots,x_m) \in \mathbb{N}^{m+1}$ such that $x_0 = x$, $x_m = x'$ and for all $i \leq m-1$, $|x_{i+1} - x_i| \leq 1$. Similarly, we let $\mathbb{T}_m^{>1}(x)$ be the set of the labeled trees $(T,l) \in \mathbb{T}_m^+(x)$ such that $l > 1$, and $\mathcal{M}_{m,x\to x'}^{>1}$ be the set of the walks $(x_1,\dots,x_m) \in \mathcal{M}_{m,x\to x'}^+$ such that $x_0,\dots,x_m > 1$. Also recall that $\mu_{(x_0,\dots,x_m)}$ denotes the distribution of a "uniform infinite" forest with root labels $x_0,\dots,x_m$.

For all $a,b,c,k' \geq 1$, $\underline{x} \in \mathcal{M}_{a,1\to k'}^+$, $\underline{y} \in \mathcal{M}_{b,2\to k'}^{>1}$, $\underline{z} \in \mathcal{M}_{c+1,k'\to k}^{>1}$, let $A_\infty^{(k)}(a,b,c,\underline{x},\underline{y},\underline{z})$ denote the event:

$$a_\infty^{(k)} = a, b_\infty^{(k)} = b, c_\infty^{(k)} = c, \quad (X_{\infty,1}^{(k)},\dots,X_{\infty,a}^{(k)}) = \underline{x},$$
$$(Y_{\infty,1}^{(k)},\dots,Y_{\infty,b}^{(k)}) = \underline{y}, \qquad (Z_{\infty,0}^{(k)},\dots,Z_{\infty,c}^{(k)}) = \underline{z}.$$

Corollary 3.2. *For all $a, b, c, k' \geq 1$, $\underline{x} \in \mathcal{M}^+_{a,1\to k'}$, $\underline{y} \in \mathcal{M}^{>1}_{b,2\to k'}$, $\underline{z} \in \mathcal{M}^{>1}_{c+1,k'\to k}$, we have*

$$\mathbb{P}\left(A^{(k)}_\infty(a,b,c,\underline{x},\underline{y},\underline{z})\right) = \frac{a+b+1}{3^{a+b+c}}\left(\prod_{i=1}^{a} w(x_i)\right)\left(\prod_{i=1}^{b} w(y_i-1)\right)\left(\prod_{i=1}^{c+1} w(z_i)w(z_i-1)\right).$$

Moreover, conditionally on $A^{(k)}_\infty(a,b,c,\underline{x},\underline{y},\underline{z})$, with the conventions $x_0 = y_0 = 0$:

— *The forests $(\tau_{\infty,i})_{1\leq i\leq a+c}$, $(\tau'_{\infty,i})_{1\leq i\leq b+c}$ and $(\overline{\tau}_{\infty,0},\ldots,\overline{\tau}_{\infty,a},\overline{\tau}'_{\infty,0},\ldots,\overline{\tau}'_{\infty,b-1})$ are independent.*

— *The trees $\tau_{\infty,i}$, $1 \leq i \leq a+c$ are independent random variables, respectively distributed according to $\rho^+_{(x_i)}$, $1 \leq i \leq a$ and $\rho^+_{(z_{a+i})}$, $a+1 \leq i \leq a+c$.*

— *The trees $\tau'_{\infty,i}$, $1 \leq i \leq b+c$ are independent random variables, obtained by adding 1 to the labels of trees distributed according to $\rho^+_{(y_i-1)}$, $1 \leq i \leq b$ and $\rho^+_{(z_{b+i})}$, $b+1 \leq i \leq a+c$, respectively.*

— *The forest $(\overline{\tau}_{\infty,0},\ldots,\overline{\tau}_{\infty,a},\overline{\tau}'_{\infty,0},\ldots,\overline{\tau}'_{\infty,b-1})$ follows the distribution $\mu_{(0,x_1,\ldots,x_a,0,y_1,\ldots,y_{b-1})}$.*

Proof. We have

$$\begin{aligned}\mathbb{P}\left(A^{(k)}_\infty(a,b,c,\underline{x},\underline{y},\underline{z})\right) = \frac{a+b+1}{6^{a+b}12^c}&\left(\prod_{i=1}^{a}\sum_{n_i\geq 0}\frac{1}{12^{n_i}}\#\mathbb{T}^+_{n_i}(x_i)\right)\left(\prod_{i=1}^{b}\sum_{n_i\geq 0}\frac{1}{12^{n_i}}\#\mathbb{T}^{>1}_{n_i}(y_i)\right)\\ &\left(\prod_{i=1}^{c+1}\sum_{n_i\geq 0}\frac{1}{12^{n_i}}\#\mathbb{T}^+_{n_i}(z_i)\right)\left(\prod_{i=1}^{c+1}\sum_{n_i\geq 0}\frac{1}{12^{n_i}}\#\mathbb{T}^{>1}_{n_i}(z_i)\right).\end{aligned}$$

Using equation (4.4) and the fact that $\#\mathbb{T}^{>1}_m(x) = \#\mathbb{T}^+_m(x-1)$ for all $m, x \in \mathbb{N}$, this gives

$$\begin{aligned}\mathbb{P}\left(A^{(k)}_\infty(a,b,c,\underline{x},\underline{y},\underline{z})\right) &= \frac{a+b+1}{6^{a+b}12^c}\left(\prod_{i=1}^{a} 2w(x_i)\right)\left(\prod_{i=1}^{b} 2w(y_i-1)\right)\left(\prod_{i=1}^{c+1} 4w(z_i)w(z_i-1)\right)\\ &= \frac{a+b+1}{3^{a+b+c}}\left(\prod_{i=1}^{a} w(x_i)\right)\left(\prod_{i=1}^{b} w(y_i-1)\right)\left(\prod_{i=1}^{c+1} w(z_i)w(z_i-1)\right),\end{aligned}$$

hence the first part of the Lemma. The conditional distributions of the trees $\tau_{\infty,i}$, $\tau'_{\infty,i}$, $\overline{\tau}_{\infty,i}$ and $\overline{\tau}'_{\infty,i}$ are then obtained exactly as in the proof of Proposition 2.1. □

3.2 Proof of the joint convergence

As in Section 2, the main step of the proof of the convergence is to show the convergence of the labels on the branches $\mathrm{x}^{(k)}_{\infty,i}$, $i \geq 1$ and $\mathrm{y}^{(k)}_{\infty,i}$, $i \geq 1$. Fix $r \in \mathbb{N}$. For all $k \in \mathbb{N}$, and for all continuous bounded functions F, G from $\mathbb{R}^r$ into $\mathbb{R}$, we let

$$\mathcal{E}_k(F,G) := \mathbb{E}\left[F(X^{(k)}_{\infty,1},\ldots,X^{(k)}_{\infty,r})G(Y^{(k)}_{\infty,1},\ldots,Y^{(k)}_{\infty,r})\mathbf{1}_{\{a^{(k)}_\infty,b^{(k)}_\infty\geq r\}}\right].$$

Lemma 3.3. *We have the convergence*

$$\mathcal{E}_k(F,G) \xrightarrow[k\to\infty]{} \mathbb{E}_1[F(\tilde{X}_0,\ldots,\tilde{X}_{r-1})]\mathbb{E}_1[G(\tilde{X}_0+1,\ldots,\tilde{X}_{r-1}+1)].$$

Proof. As in the previous section, we introduce independent random walks $\hat{X}$, $\hat{Y}$ and $\hat{Z}$ with uniform steps in $\{-1,0,1\}$, and consider associated martingales MX, MY and MZ such that for all $j \geq 0$,

$$MX_j = \frac{f(\hat{X}_j)}{f(\hat{X}_0)}\prod_{i=0}^{j-1} w(\hat{X}_i) \qquad MY_j = \frac{f(\hat{Y}_j)}{f(\hat{Y}_0)}\prod_{i=0}^{j-1} w(\hat{Y}_i) \qquad MZ_j = \frac{g(\hat{Z}_j)}{g(\hat{Z}_0)}\prod_{i=0}^{j-1} v(\hat{Z}_i),$$

where $v(x) = w(x)w(x+1) = x(x+4)/(x+2)^2$ and $g(x) = x(x+4)(5x^2+20x+17)$ for all $x \in \mathbb{N}$. From now on, we work under the assumption $1 \le r \le k$. With the above notation, we can write

$$\begin{aligned}\mathcal{E}_k(F,G) = \sum_{a,b\ge r-1}\sum_{c\ge 0}\frac{a+b+1}{9}\sum_{k'\ge 1}\mathbb{E}\Bigg[&\mathbb{E}_1\left[F(\hat{X}_0,\dots,\hat{X}_{r-1})\frac{f(1)w(k'+1)}{f(k'+1)}MX_{a-1}\mathbf{1}_{\{\hat{X}_{a-1}=k'+1\}}\right]\\ &\mathbb{E}_1\left[G(\hat{Y}_0+1,\dots,\hat{Y}_{r-1}+1)\frac{f(1)w(k')}{f(k')}MY_{b-1}\mathbf{1}_{\{\hat{Y}_{b-1}=k'\}}\right]\\ &\mathbb{E}_{k'}\left[\frac{g(k')v(k-1)}{g(k-1)}MZ_c\mathbf{1}_{\{\hat{Z}_c=k-1\}}\right]\Bigg].\end{aligned}$$

Using the Markov property and re-arranging the terms yields

$$\begin{aligned}\mathcal{E}_k(F,G) = \sum_{k'\ge 1}\mathbb{E}\Bigg[&MX_{r-1}F(\hat{X}_0,\dots,\hat{X}_{r-1})MY_{r-1}G(\hat{Y}_0+1,\dots,\hat{Y}_{r-1}+1)\frac{f(1)w(k'+1)}{3f(k'+1)}\\ &\frac{f(1)w(k')}{3f(k')}\sum_{a,b\ge 0}(a+b+2r-1)\mathbb{E}_{\hat{X}_{r-1}}\left[MX'_a\mathbf{1}_{\{\hat{X}'_a=k'+1\}}\right]\mathbb{E}_{\hat{Y}_{r-1}}\left[MY'_b\mathbf{1}_{\{\hat{Y}'_b=k'\}}\right]\\ &\frac{g(k')v(k-1)}{g(k-1)}\sum_{c\ge 0}\mathbb{E}_{k'}\left[MZ_c\mathbf{1}_{\{\hat{Z}_c=k-1\}}\right]\Bigg],\end{aligned}$$

where $\hat{X}', \hat{Y}', MX', MY'$ are independent copies of $\hat{X}, \hat{Y}, MX, MY$. We already have the necessary ingredients in Section 2 to study the first factors; the only additional quantity we need to compute is

$$H'_{k'}(k) = \sum_{c\ge 0}\mathbb{E}_{k'}\left[MZ_c\mathbf{1}_{\{\hat{Z}_c=k\}}\right] = \sum_{c\ge 0}\mathbb{P}_{k'}\left(\tilde{\hat{Z}}_c = k\right),$$

where $\tilde{\hat{Z}}$ is the image of $\hat{Z}$ under the measure-change given by the martingale MZ, i.e. the Markov process such that $\mathbb{E}[\phi(\tilde{\hat{Z}}_i)] = \mathbb{E}[MZ_i\phi(\hat{Z}_i)]$ for every continuous bounded function ϕ.

Lemma 3.4. *Fix $k, k' \ge 2$. We have the following equalities:*

— *if $k' \le k$,*

$$\frac{g(k')v(k)}{g(k)}H'_{k'}(k) = \frac{3g(k')}{35(k+1)(k+2)(k+3)}$$

— *if $k' > k$,*

$$\frac{g(k')v(k)}{g(k)}H'_{k'}(k) = \frac{3g(k)}{35(k'+1)(k'+2)(k'+3)}.$$

We omit the technical detail of the proof of this result; the ideas are exactly the same as in the proof of Lemma 2.2. Now

$$\begin{aligned}\mathcal{E}_k(F,G) = \sum_{1\le x,y\le r}\left(\sum_{k'\ge 1}\mathcal{H}_{x,y,k'}(k)\right)&\mathbb{E}_1[MX_{r-1}F(\hat{X}_0,\dots,\hat{X}_{r-1})\mathbf{1}_{\{\hat{X}_{r-1}=x\}}]\\ &\mathbb{E}_1[MY_{r-1}G(\hat{Y}_0+1,\dots,\hat{Y}_{r-1}+1)\mathbf{1}_{\{\hat{Y}_{r-1}=y\}}],\end{aligned}$$

where

$$\begin{aligned}\mathcal{H}_{x,y,k'}(k) =&\frac{f(1)w(k'+1)}{3f(k'+1)}\frac{f(1)w(k')}{3f(k')}\frac{g(k')v(k-1)}{g(k-1)}H'_{k'}(k-1)\\ &\times(H^*_x(k'+1)H_y(k') + H_x(k'+1)H^*_y(k') + (2r-5)H_x(k'+1)H_y(k')).\end{aligned}$$

Therefore, it is enough to show that $\sum_{k'\geq 1}\mathcal{H}_{x,y,k'}(k)$ converges to 1 as $k\to\infty$, uniformly in $x,y\leq r$.

Let us first treat the terms for which $k'\geq k$. We have

$$\frac{g(k')v(k-1)}{g(k-1)}H'_{k'}(k-1)=\frac{3g(k-1)}{35(k'+1)(k'+2)(k'+3)}.$$

Moreover, the results of Lemmas 2.2 and 2.3 show that, uniformly in $y\leq r$,

$$\frac{f(1)w(k')}{3f(k')}H^*_y(k')\xrightarrow[k'\to\infty]{}1$$
$$\frac{f(1)w(k')}{3f(k')}H_y(k')\sim_{k'\to\infty}\frac{2}{(k')^2},$$

and that the same holds with $k'+1$ instead of k' in the left-hand term. As a consequence, we have

$$\begin{aligned}\sum_{k'\geq k}\mathcal{H}_{x,y,k'}(k)&\sim_{k\to\infty}\frac{3g(k-1)}{35}\sum_{k'\geq k}\frac{4}{(k')^5}\\&\sim_{k\to\infty}\frac{3k^4}{7}\frac{1}{k^4}=\frac{3}{7},\end{aligned}\tag{4.12}$$

uniformly in $x,y\leq r$.

In second, we consider the terms for which we have $x\vee y<k'\leq k-1$. Lemmas 2.2 and 2.3 yield the following estimates, uniformly in $y\leq r$:

$$\frac{f(1)w(k')}{3f(k')}H^*_y(k')=1-\frac{3}{10(k'+1)(k'+2)}\left(C_y+1-\left(1\wedge\frac{h(k')}{h(y)}\right)\right)$$

and

$$\frac{f(1)w(k')}{3f(k')}H_y(k')\sim_{k'\to\infty}\frac{2}{(k')^2}.$$

Putting this together with the result of Lemma 3.4, we get

$$\begin{aligned}\sum_{k'=x\vee y+1}^{k-2}\mathcal{H}_{x,y,k'}(k)&\sim_{k\to\infty}\frac{3}{35k^3}\sum_{k'=x\vee y+1}^{k-1}2\times\frac{2}{(k')^2}\times 5(k')^4\\&\sim_{k\to\infty}\frac{12}{7k^3}\frac{k^3}{3}=\frac{4}{7}.\end{aligned}\tag{4.13}$$

The remaining term is

$$\sum_{k'=1}^{x\vee y}\mathcal{H}_{x,y,k'}(k)=O\left(\frac{1}{k^3}\right).$$

Putting this together with (4.12) and (4.13), we obtain

$$\sum_{k'\geq 1}\mathcal{H}_{x,y,k'}(k)\xrightarrow[k\to\infty]{}\frac{3}{7}+\frac{4}{7}=1,$$

uniformly in $x,y\leq r$, hence the conclusion. □

To complete the proof of Theorem 1.4, we finally come back to the trees attached on the branches $\mathrm{x}^{(k)}_{\infty,i}$, $i\geq 1$ and $\mathrm{y}^{(k)}_{\infty,i}$, $i\geq 1$, putting together the above result and Corollary 3.2. Let $E^{(k)}(r)$ be the event that $a^{(k)}_\infty,b^{(k)}_\infty\geq r$, and the trees $(\overline{\tau}_{\infty,i})^{(k)}$ and $(\overline{\tau}'_{\infty,i})^{(k)}$ are finite for all $i\leq r$. Conditionally on $E^{(k)}(r)$, we have the following properties on the spines of $\theta^{(k)}_\infty$ and $\theta^{(-k+1)}_\infty$:

- The points $\mathfrak{s}_i(\theta_\infty^{(k)})$ and $\mathrm{x}_{\infty,i}^{(k)}$ are the same for all $i \le r$, hence $R_i(\theta_\infty^{(k)}) = \tau_{\infty,i}^{(k)}$ and $L_i(\theta_\infty^{(k)}) = \overline{\tau}_{\infty,i}^{(k)}$ for all $i < r$.
- The points $\mathfrak{s}_i(\theta_\infty^{(-k+1)})$ and $\mathrm{y}_{\infty,i}^{(k)}$ are the same for all $i \le r$, hence $R_i(\theta_\infty^{(-k+1)}) = (\overline{\tau}'_{\infty,i})^{(k)}$ and $L_i(\theta_\infty^{(-k+1)}) = (\tau'_{\infty,i})^{(k)}$ for all $i < r$.
- As a consequence, the spine labels $(S_i(\theta_\infty^{(k)}), S_i(\theta_\infty^{(-k+1)}) - 1)_{1 \le i \le r}$ converge in distribution to $(\tilde{X}_i, \tilde{Y}_i)_{0 \le i \le r-1}$, with $\tilde{X}_0 = \tilde{Y}_0 = 1$.

Further conditioning on $(S_i(\theta_\infty^{(k)}), S_i(\theta_\infty^{(-k+1)}))_{0 \le i < r}$, we get that:

- The subtrees $L_i(\theta_\infty^{(k)})$, $0 \le i < r$ and $R_i(\theta_\infty^{(k)})$, $1 \le i < r$ are independent random variables, with respective distributions $\rho_{(S_i(\theta_\infty^{(k)}))}$ and $\rho^+_{(S_i(\theta_\infty^{(k)}))}$.
- The subtrees $L_i(\theta_\infty^{(-k+1)})$, $0 \le i < r$ and $R_i(\theta_\infty^{(-k+1)})$, $1 \le i < r$ are independent random variables, respectively obtained by adding 1 to the labels of trees distributed according to $\rho^+_{(S_i(\theta_\infty^{(k)})-1)}$ and $\rho_{(S_i(\theta_\infty^{(k)})-1)}$.
- The random forests $(L_i(\theta_\infty^{(k)}, R_i(\theta_\infty^{(k)}))_{0 \le i < r}$ and $(L_i(\theta_\infty^{(-k+1)}, R_i(\theta_\infty^{(-k+1)}))_{0 \le i < r}$ are independent.

Therefore, it is enough to show that $\mathbb{P}(\overline{E^{(k)}(r)})$ converges to 0 as $k \to \infty$. Fix $\varepsilon > 0$. We have

$$\mathbb{P}(\overline{E^{(k)}(r)}) \le \mathbb{P}(a_\infty^{(k)} < r \text{ or } b_\infty^{(k)} < r) + \mathbb{E}\left[1 \wedge \frac{2r+2}{a_\infty^{(k)} + b_\infty^{(k)} + 1}\right]$$

We know from Lemma 3.3 that the first term converges to 0. More precisely, for all $r' \in \mathbb{N}$, we have

$$\mathbb{P}(a_\infty^{(k)} < r' \text{ or } b_\infty^{(k)} < r') \le \varepsilon$$

for all k large enough, hence

$$\mathbb{E}\left[1 \wedge \frac{2r+2}{a_\infty^{(k)} + b_\infty^{(k)} + 1}\right] \le \varepsilon + \frac{2r+2}{2r'+1}$$

for k large enough. Thus we can choose r' in such a way that for all k large enough, we have

$$\mathbb{P}(\overline{E^{(k)}(r)}) \le 3\varepsilon.$$

This concludes the proof.

4 Convergence of the associated quadrangulations

As indicated in the Introduction, the main step of the proof of Theorem 1.5 consists in showing the following result. We use the conventions

$$\theta_\infty^{(\infty)} = \overrightarrow{\theta}_\infty, \qquad \theta_\infty^{(-\infty)} = \overleftarrow{\theta}_\infty,$$
$$\overrightarrow{Q}_\infty^{(\infty)} = \overrightarrow{Q}_\infty, \qquad \overleftarrow{Q}_\infty^{(\infty)} = \overleftarrow{Q}_\infty.$$

Proposition 4.1. *For all $r \in \mathbb{N}$ and $\varepsilon > 0$, there exists $h \in \mathbb{N}$ such that for all k large enough, possibly infinite, we have*

$$V\left(B_{\overrightarrow{Q}_\infty^{(k)}}(r)\right) \subset V\left(B_{\theta_\infty^{(k)}}(h)\right) \tag{4.14}$$

and

$$V\left(B_{\overleftarrow{Q}_\infty^{(k)}}(r)\right) \subset V\left(B_{\theta_\infty^{(-k+1)}}(h)\right) \cup \{\lambda_i : |i| \le r\} \tag{4.15}$$

with probability at least $1 - \varepsilon$.

Let us first see how this result allows us to prove the theorem.

Proof of Theorem 1.5. Using the Skorokhod representation theorem, we assume that the convergence

$$(\theta_\infty^{(k)}, \theta_\infty^{(-k+1)}) \xrightarrow[k\to\infty]{} (\overrightarrow{\theta_\infty}, \overleftarrow{\theta_\infty}),$$

obtained in Theorem 1.4, holds almost surely. In particular, it also holds in probability: for all $h \in \mathbb{N}$ and $\varepsilon > 0$, we have

$$\mathbb{P}\left(D(\theta_\infty^{(k)}, \overrightarrow{\theta_\infty}) \le \frac{1}{1+h} \text{ and } D(\theta_\infty^{(-k+1)}, \overleftarrow{\theta_\infty}) \le \frac{1}{1+h} \right) \ge 1-\varepsilon$$

for all k large enough, which means that

$$B_{\theta_\infty^{(k)}}(h) = B_{\overrightarrow{\theta_\infty}}(h) \qquad \text{and} \qquad B_{\theta_\infty^{(-k)}}(h) = B_{\overleftarrow{\theta_\infty}}(h) \tag{4.16}$$

with probability at least $1-\varepsilon$, for all k large enough.

For all $r \in \mathbb{N}$ and $\varepsilon > 0$, the above proposition shows that there exists h_ε such that the inclusions (4.14) and (4.15) hold with probability at least $1-\varepsilon$, for all k large enough. Putting this together with (4.16) for $h = h_\varepsilon$, we get that

$$B_{\overrightarrow{Q}_\infty^{(k)}}(r) = B_{\Phi(\overrightarrow{\theta_\infty})}(r) \qquad \text{and} \qquad B_{\overleftarrow{Q}_\infty^{(k)}}(r) = B_{\Phi(\overleftarrow{\theta_\infty})}(r)$$

with probability at least $1-\varepsilon$, for all k large enough (possibly infinite). Therefore, we have the convergence

$$(\overrightarrow{Q}_\infty^{(k)}, \overleftarrow{Q}_\infty^{(k)}) \xrightarrow[k\to\infty]{} (\overrightarrow{Q}_\infty, \overleftarrow{Q}_\infty)$$

in probability, hence the joint distributional convergence. □

The rest of the section is devoted to the proof of Proposition 4.1. We first introduce conditions on the "left-hand side" and "right-hand side" of the trees $\theta_\infty^{(k)}$, $\theta_\infty^{(-k+1)}$, which are sufficient to get the ball inclusions (4.14) and (4.15). This is done in Section 4.1 (see in particular Lemma 4.3). In Sections 4.2, 4.3 and 4.4, we then show that an "elementary block" of these conditions holds with arbitrarily high probability, for all s and k large enough. The corresponding results are stated in Lemmas 4.4 and 4.5. Finally, Section 4.5 concludes the proof of the proposition.

4.1 Conditions on the right-hand and left-hand part of a labeled tree

We first introduce some more detailed notation for the balls in a rooted tree T. For all $s \ge 0$, we let $\partial B_T(s)$ denote the "boundary" of the ball of radius s, defined as

$$\partial B_T(s) = \{ \mathrm{v} \in T : \mathrm{v} \text{ has height } s \}.$$

In what follows, the letter L will correspond to the "left-hand part" of a tree, and R will be used for the "right-hand part". All the following notations are given for the left-hand part, and are also valid for the right-hand part (replacing L by R). Assume that $T \in \mathbf{S}$, and recall that $L_i(T)$ denotes the subtree of the descendants of $\mathfrak{s}_i(T)$ that are on the left of the spine. We let

$$L(T) = \bigcup_{i \ge 0} L_i(T),$$

and for all $s \ge 0$,

$$LB_T(s) = B_T(s) \cap L(T) = \bigcup_{i=0}^{s} B_{L_i(T)}(s-i),$$

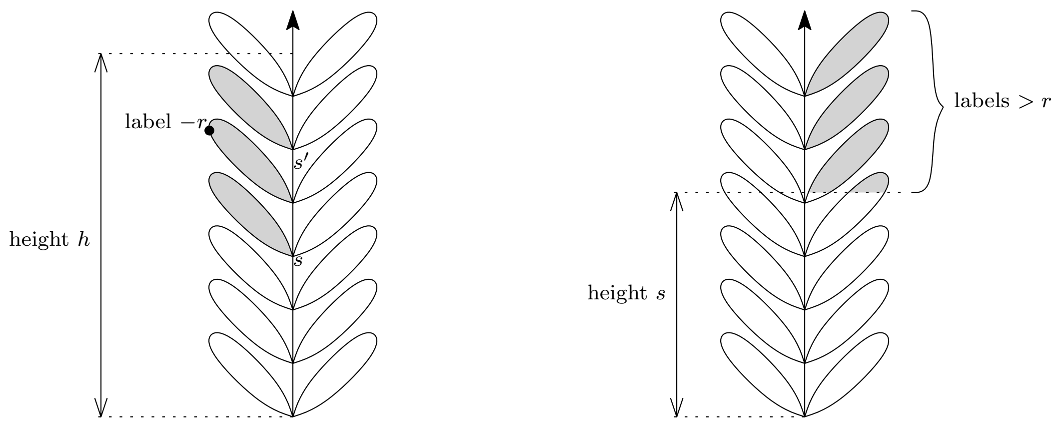

Figure 4.6 – An illustration of the conditions $\theta \in \mathbb{A}_L(r,s,s',h)$ (on the left) and $\theta \in \mathbb{A}_{R+}(r,s)$ (on the right).

and

$$\partial LB_T(s) = \partial B_T(s) \cap L(T).$$

We also use the natural extensions of this notation to labeled trees.

We are interested in the following subsets of $\mathbb{S}$, for all $r,s,s',h \in \mathbb{N}$:

$$\mathbb{A}_L(r,s,s',h) = \left\{ (T,l) \in \mathbb{S} : \bigcup_{i=0}^{s'} L_i(T) \subset B_T(h), \text{ and } \exists \mathrm{v} \in \bigcup_{i=s+1}^{s'} L_i(T) \text{ s.t. } l(\mathrm{v}) = -r \right\}$$

and

$$\mathbb{A}_{L+}(r,s) = \{(T,l) \in \mathbb{S} : \forall \mathrm{v} \in L(T) \setminus LB_T(s),\ l(\mathrm{v}) > r\}.$$

Figure 4.6 illustrates these definitions. We give a sufficient condition for an inclusion between the balls in θ and in $\Phi(\theta)$, in terms of these sets $\mathbb{A}_L(r,s,s',h)$, $\mathbb{A}_{L+}(r,s)$, $\mathbb{A}_R(r,s,s',h)$ and $\mathbb{A}_{R+}(r,s)$:

Lemma 4.2. *Let $r \in \mathbb{N}$.*

1. *For all $\theta \in \overrightarrow{\mathbb{S}}$, if there exists sequences $(s(r'))_{0 \le r' \le r}$ and $(h(r'))_{1 \le r' \le r}$ such that*

$$\theta \in \mathbb{A}_L(r', s(r'-1), s(r'), h(r')) \cap \mathbb{A}_{R+}(r', s(r')) \qquad \forall r' \in \{1,\dots,r\},$$

then we have $V(B_{\Phi(\theta)}(r)) \subset V(B_\theta(h(r)))$.

2. *For all $\theta \in \overleftarrow{\mathbb{S}}$, if there exists sequences $(s(r'))_{0 \le r' \le r}$ and $(h(r'))_{1 \le r' \le r}$ such that*

$$\theta \in \mathbb{A}_R(r', s(r'-1), s(r'), h(r')) \cap \mathbb{A}_{L+}(r', s(r')) \qquad \forall r' \in \{1,\dots,r\},$$

then we have $V(B_{\Phi(\theta)}(r)) \subset V(B_\theta(h(r))) \cup \{\lambda_i : |i| \le r\}$.

Proof. Let $\theta \in \overrightarrow{\mathbb{S}}$. We show by induction that for all $r \ge 0$, if there exists sequences $(s(r'))_{0 \le r' \le r}$ and $(h(r'))_{1 \le r' \le r}$ such that

$$\theta \in \mathbb{A}_L(r, s(r'-1), s(r'), h(r')) \cap \mathbb{A}_{R+}(r', s(r')) \qquad \forall r' \in \{1,\dots,r\},$$

then we have

$$V\left(B_{\Phi(\theta)}(r)\right) \subset V\left(RB_\theta(s(r)) \cup \bigcup_{i=0}^{s(r)} L_i(\theta) \right).$$

This is enough to prove the first part of the Lemma. Indeed, since θ belongs to $\mathbb{A}_L(r, s(r-1), s(r), h(r))$, we have $\bigcup_{i=0}^{s(r)} L_i(\theta) \subset B_\theta(h(r))$ and $s(r) \leq h(r)$, so

$$V\left(RB_\theta(s(r)) \cup \bigcup_{i=0}^{s(r)} L_i(\theta)\right) \subset V(B_\theta(h(r))).$$

The result is obviously true for $r = 0$. Assume that it holds for a given $r \geq 0$. We order the corners of θ by writing $\mathrm{c}_n(\theta) \leq \mathrm{c}_{n'}(\theta)$ for all $n \leq n'$. For all $r' \leq r+1$, let $\xi_{r'}$ denote the largest corner incident to the vertex $\mathfrak{s}_{s(r')}$. Note that for all $r' \leq r$, for every corner c of θ, we have $\mathrm{c} \leq \xi_{r'}$ if and only if every corner $\tilde{\mathrm{c}}$ incident to the same vertex as c verifies $\tilde{\mathrm{c}} \leq \xi_{r'}$. The induction hypothesis ensures that for every corner c of θ which is incident to a vertex of $B_{\Phi(\theta)}(r)$, we have $\mathrm{c} \leq \xi_r$. (This is the case even if the corresponding vertex is in the right-hand part of θ.)

Let $\mathrm{v} \in V(\theta)$. The vertex v belongs to $B_{\Phi(\theta)}(r+1)$ if and only if one of the following conditions holds:

1. v belongs to $B_{\Phi(\theta)}(r)$.
2. There exist a vertex v' of $B_{\Phi(\theta)}(r)$, and two corners c and c', respectively incident to v and v', such that $\sigma_\theta(\mathrm{c}) = \mathrm{c}'$.
3. There exist a vertex v' of $B_{\Phi(\theta)}(r)$, and two corners c and c', respectively incident to v and v', such that $\sigma_\theta(\mathrm{c}') = \mathrm{c}$.

Respectively, in these three cases, it holds that:

1. Every corner $\tilde{\mathrm{c}}$ incident to v is such that $\tilde{\mathrm{c}} \leq \xi_r \leq \xi_{r+1}$.
2. We have $\mathrm{c} \leq \mathrm{c}' \leq \xi_r$, so every corner $\tilde{\mathrm{c}}$ incident to v is such that $\tilde{\mathrm{c}} \leq \xi_r \leq \xi_{r+1}$.
3. The corner c is the first corner with label $l(\mathrm{v}') - 1$ after c'. Since v' belongs to $B_{\Phi(\theta)}(r)$, the bound (4.1) ensures that
$$d_{\Phi(\theta)}(\mathrm{v}_0, \mathrm{v}') \geq \left| l(\mathrm{v}_0) - l(\mathrm{v}') \right| = l(\mathrm{v}')$$
(where v_0 denotes the root of θ), so $l(\mathrm{v}') - 1 \geq -r-1$. Moreover, we have $\mathrm{c}' \leq \xi_r$, and since θ belongs to $\mathbb{A}_L(r+1, s(r), s(r+1), h(r+1))$, there exists a corner with label $-r-1$ between ξ_r and ξ_{r+1}. As a consequence, we have $\mathrm{c} \leq \xi_{r+1}$, and therefore every corner $\tilde{\mathrm{c}}$ incident to v is such that $\tilde{\mathrm{c}} \leq \xi_{r+1}$.

Thus, we get the inclusion

$$V\left(B_{\Phi(\theta)}(r+1)\right) \subset V\left(R(\theta) \cup \bigcup_{i=0}^{s(r+1)} L_i(\theta)\right).$$

Finally, for every vertex $\mathrm{v} \in R(\theta) \setminus RB_\theta(s(r+1))$, since θ belongs to $\mathbb{A}_{R+}(r+1, s(r+1))$, we have $l(\mathrm{v}) > r+1$, so v is at distance at least $r+2$ of the root in $\Phi(\theta)$. This yields

$$V\left(B_{\Phi(\theta)}(r+1)\right) \subset V\left(RB_\theta(s(r+1)) \cup \bigcup_{i=0}^{s(r+1)} L_i(\theta)\right).$$

We now consider the case where $\theta \in \overleftarrow{\mathbb{S}}$. Similarly, it is enough to show by induction that for all $r \geq 0$, if there exists sequences $(s(r'))_{0 \leq r' \leq r}$ and $(h(r'))_{1 \leq r' \leq r}$ verifying the hypotheses, then we have

$$V\left(B_{\Phi(\theta)}(r) \setminus \Lambda\right) \subset V\left(LB_\theta(s(r)) \cup \bigcup_{i=0}^{s(r)} R_i(\theta)\right).$$

(Indeed, equation (4.1) shows that $V(B_{\Phi(\theta)}(r) \cap \Lambda) \subset \{\lambda_i : |i| \leq r\}$.) Assume that the result holds for a given $r \geq 0$. For all $r' \leq r+1$, let $\xi'_{r'}$ denote the smallest corner incident to the vertex $\mathfrak{s}_{s(r')}$. For every corner c of θ which is incident to a vertex of $B_{\Phi(\theta)}(r)$, we have $\mathrm{c} \geq \xi_r$. We fix $\mathrm{v} \in V(\theta)$, and study the same three cases as above. Respectively, we obtain that:

1. Every corner $\tilde{\mathrm{c}}$ incident to v is such that $\tilde{\mathrm{c}} \geq \xi'_r \geq \xi'_{r+1}$.
2. The corner c' is the first corner with label $l(\mathrm{v})-1$ after c (or a point of Λ, if such a corner does not exist), and equation (4.1) gives that $l(\mathrm{v}) - 1 = l(\mathrm{v}') \geq -r$. Since θ belongs to $\mathbb{A}_R(r+1, s(r), s(r+1), h(r+1))$, there exists a corner with label $-r-1$ which is (strictly) between ξ'_{r+1} and ξ'_r. So, if we had $\mathrm{c} < \xi'_{r+1}$, this would imply $\mathrm{c}' < \xi'_r$, which is impossible since v' is in $B_{\Phi(\theta)}(r)$. Thus, we have $\mathrm{c} \geq \xi'_{r+1}$, and every corner $\tilde{\mathrm{c}}$ incident to v is such that $\tilde{\mathrm{c}} \geq \xi'_{r+1}$.
3. Note that since v is a vertex of θ, we cannot have $\mathrm{v}' \in \Lambda$. Thus, we have $\mathrm{c} \geq \mathrm{c}' \geq \xi'_r$, so every corner $\tilde{\mathrm{c}}$ incident to v is such that $\tilde{\mathrm{c}} \geq \xi'_r \geq \xi'_{r+1}$.

This yields the inclusion

$$V\left(B_{\Phi(\theta)}(r+1) \setminus \Lambda\right) \subset V\left(L(\theta) \cup \bigcup_{i=0}^{s(r+1)} R_i(\theta)\right),$$

and the same argument as above concludes the proof. □

Our goal is now to obtain similar conditions on the trees $\theta_\infty^{(k)}$ and $\theta_\infty^{(-k+1)}$, sufficient to get the ball inclusions (4.14) and (4.15). Note that we cannot apply the above result directly, since $\theta_\infty^{(k)}$ and $\theta_\infty^{(-k+1)}$ are elements of $\mathbb{S}^*(0)$ and $\mathbb{S}^*(1)$ instead of $\overrightarrow{\mathbb{S}}$ and $\overleftarrow{\mathbb{S}}$. Moreover, for example in $\theta_\infty^{(k)}$, we are not interested in *all* the vertices which are on the right of the spine, but only in those which are on the right of the segment $[\![\mathrm{e}_k(\theta_\infty), \mathrm{e}_0(\theta_\infty)]\!]$. Informally, the others are "cut-off" from the root when we split the quadrangulation Q_∞ along the maximal geodesic, so they do not belong to the neighbourhood of $\mathrm{e}_k(\theta_\infty)$ in $\overrightarrow{Q}_\infty^{(k)}$.

Therefore, for all $k \in \mathbb{N}$, we further decompose the trees $\theta_\infty^{(k)}$ and $\theta_\infty^{(-k+1)}$. Recall the notation introduced in Section 3.1. We let

$$R_\infty^{(k)} = \bigcup_{i=1}^{a_\infty^{(k)}+c_\infty^{(k)}} \tau_{\infty,i}^{(k)} \qquad \text{and} \qquad R_\infty^{(k)}(s) = R_\infty^{(k)} \cap B_{\theta_\infty^{(k)}}(s) \quad \forall s \geq 0,$$

and similarly,

$$L_\infty^{(-k+1)} = \bigcup_{i=1}^{b_\infty^{(k)}+c_\infty^{(k)}} (\tau'_{\infty,i})^{(k)} \qquad \text{and} \qquad L_\infty^{(-k+1)}(s) = L_\infty^{(-k+1)} \cap L_{\theta_\infty^{(-k+1)}}(s) \quad \forall s \geq 0.$$

Note that we have, for example, $R_\infty^{(k)} \subset R(\theta_\infty^{(k)})$ and $R_\infty^{(k)}(s) \subset RB_{\theta_\infty^{(k)}}(s)$. We consider the following events:

— $\mathcal{A}_{R+}^{(k)}(r,s)$: "every vertex $\mathrm{v} \in R_\infty^{(k)} \setminus (R_\infty^{(k)}(s))$ has label greater than r in $\theta_\infty^{(k)}$",

— $\mathcal{A}_{L+}^{(-k+1)}(r,s)$: "every vertex $\mathrm{v} \in L_\infty^{(-k+1)} \setminus (L_\infty^{(-k+1)}(s))$ has label greater than r in $\theta_\infty^{(-k+1)}$".

For $k = \infty$, we complement this notation by setting

$$\mathcal{A}_{R+}^{(\infty)}(r,s) = \left\{\overrightarrow{\theta}_\infty \in \mathbb{A}_{R+}(r,s)\right\} \quad \text{and} \quad \mathcal{A}_{L+}^{(-\infty)}(r,s) = \left\{\overleftarrow{\theta}_\infty \in \mathbb{A}_{L+}(r,s)\right\}.$$

We can now adapt Lemma 4.2 to $\theta_\infty^{(k)}$ in the following way:

Lemma 4.3. *Let $r \in \mathbb{N}$, and consider two sequences of positive integers $(s(r'))_{0 \leq r' \leq r}$ and $(h(r'))_{1 \leq r' \leq r}$. For all $k \in \mathbb{N} \cup \{\infty\}$, we have that:*

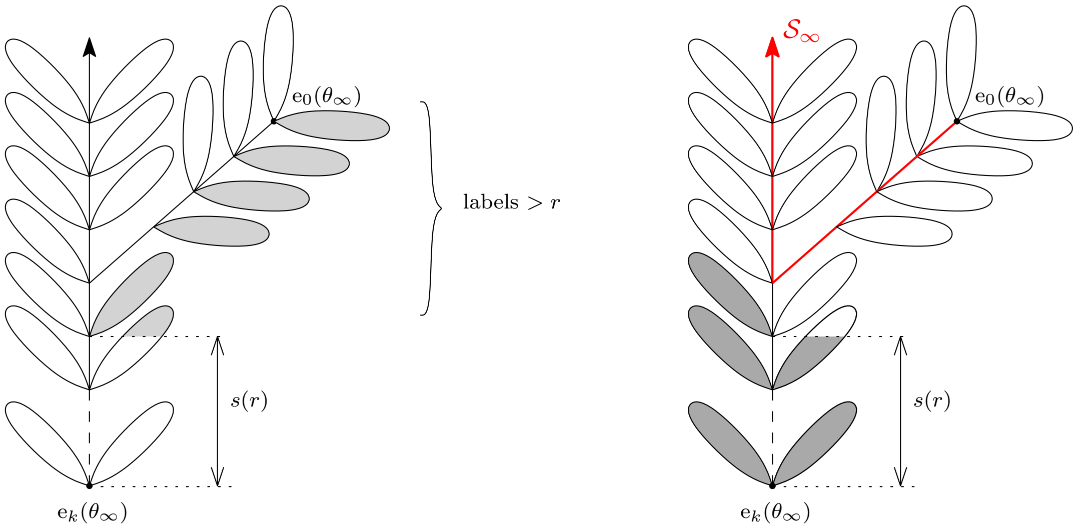

Figure 4.7 – Illustration of the event $\mathcal{A}_{R+}(s(r))$ (on the left), and of the additional condition $\mathfrak{s}_{s(r)+1}(\theta_\infty^{(k)}) \prec \mathrm{e}_0(\theta_\infty)$ in $\theta_\infty^{(k)}$ (on the right). The second figure emphasises the fact that under the condition $\mathfrak{s}_{s(r)+1}(\theta_\infty^{(k)}) \prec \mathrm{e}_0(\theta_\infty)$, the spine $\mathcal{S}_\infty$ does not intersect the set $R_\infty^{(k)}(s(r)) \cup \bigcup_{i=0}^{s(r)} L_i(\theta_\infty^{(k)})$ (and in particular, it does not contain $e_k(\theta_\infty)$). This will be used in the proof of Lemma 4.3.

1. *Conditionally on* $\mathfrak{s}_{s(r)+1}(\theta_\infty^{(k)}) \prec \mathrm{e}_0(\theta_\infty)$ *in* $\theta_\infty^{(k)}$ *and on the event*

$$\bigcap_{r'=1}^{r} \left(\theta_\infty^{(k)} \in \mathbb{A}_L(r', s(r'-1), s(r'), h(r'))\right) \cap \mathcal{A}_{R+}^{(k)}(r', s(r')), \tag{4.17}$$

we have

$$V\left(B_{\overrightarrow{Q}_\infty^{(k)}}(r)\right) \subset V\left(B_{\theta_\infty^{(k)}}(h(r))\right)$$

almost surely.

2. *Conditionally on* $\mathfrak{s}_{s(r)+1}(\theta_\infty^{(-k+1)}) \prec \mathrm{e}_0(\theta_\infty)$ *in* $\theta_\infty^{(-k+1)}$ *and on the event*

$$\bigcap_{r'=1}^{r} \left(\theta_\infty^{(-k+1)} \in \mathbb{A}_R(r', s(r'-1), s(r'), h(r'))\right) \cap \mathcal{A}_{L+}^{(-k+1)}(r', s(r')), \tag{4.18}$$

we have

$$V\left(B_{\overleftarrow{Q}_\infty^{(k)}}(r)\right) \subset V\left(B_{\theta_\infty^{(-k+1)}}(h(r))\right) \cup \{\lambda_i : |i| \leq r\}$$

almost surely.

Figure 4.7 illustrates the "new" conditions which appear, compared to the conditions of Lemma 4.2 (both are shown for the first case). Note that the condition on the left-hand side of $\theta_\infty^{(k)}$ is exactly the same as in Lemma 4.2, already illustrated in Figure 4.6.

Proof. The case where $k = \infty$ is a direct application of Lemma 4.2. From now on, we fix $k \in \mathbb{N}$.

Let $\mathcal{S}_\infty = \{\mathfrak{s}_i(\theta_\infty) : i \geq 0\}$ be the spine of θ_∞, and $\Gamma'_\infty = \left\{\mathrm{e}'_{k'} : k' \geq 1\right\}$ be the "copy" of the infinite geodesic ray we introduced in the definition of the split quadrangulation $\mathrm{Sp}(Q_\infty)$ (see for example Figure 4.2). The construction of $\mathrm{Sp}(Q_\infty)$ ensures that there are no edges between the vertices of $(\Gamma'_\infty \cup R(\theta_\infty)) \setminus \mathcal{S}_\infty$ and the vertices of $L(\theta_\infty) \setminus \mathcal{S}_\infty$. As a consequence, any geodesic from a point of $(\Gamma'_\infty \cup R(\theta_\infty)) \setminus \mathcal{S}_\infty$ to a point of $L(\theta_\infty) \setminus \mathcal{S}_\infty$ contains a vertex of $\mathcal{S}_\infty$.

Note that we have the following equalities:

$$R(\theta_\infty) = R(\theta_\infty^{(k)}) \setminus R_\infty^{(k)} = R(\theta_\infty^{(-k+1)}) \cup L_\infty^{(-k+1)}$$
$$L(\theta_\infty) = L(\theta_\infty^{(k)}) \cup R_\infty^{(k)} = L(\theta_\infty^{(-k+1)}) \setminus L_\infty^{(-k+1)}.$$

In the first case, the same induction as in the proof Lemma 4.2 shows that conditionally on (4.17), we have

$$V\left(B_{\overrightarrow{Q}_\infty^{(k)}}(r)\right) \cap V\left(L(\theta_\infty)\right) \subset V\left(R_\infty^{(k)}(s(r)) \cup \bigcup_{i=0}^{s(r)} L_i(\theta_\infty^{(k)})\right). \tag{4.19}$$

Indeed, the first step of the induction shows that there are no vertices belonging to the ball $B_{\overrightarrow{Q}_\infty^{(k)}}(r)$ after $\mathfrak{s}_{s(r)}(\theta_\infty^{(k)})$ in the clockwise order, or equivalently

$$V\left(B_{\overrightarrow{Q}_\infty^{(k)}}(r)\right) \cap V\left(L(\theta_\infty^{(k)})\right) \subset V\left(\bigcup_{i=0}^{s(r)} L_i(\theta_\infty^{(k)})\right),$$

and since the vertices in $R_\infty^{(k)} \setminus R_\infty^{(k)}(s(r))$ all have labels greater than r, we also have

$$V\left(B_{\overrightarrow{Q}_\infty^{(k)}}(r)\right) \cap V\left(R_\infty^{(k)}\right) \subset V\left(R_\infty^{(k)}(s(r))\right).$$

Noting that $L(\theta_\infty^{(k)}) \cup R_\infty^{(k)} = L(\theta_\infty)$ yields inclusion (4.19).

To conclude the proof of the first point, we only have to show that the vertices of $R(\theta_\infty) \setminus \mathcal{S}_\infty$ are at distance at least $r+1$ from $\mathrm{e}_k(\theta_\infty)$ in $\overrightarrow{Q}_\infty^{(k)}$. Let $\mathrm{v} \in V(R(\theta_\infty) \setminus \mathcal{S}_\infty)$, and let γ be a geodesic path from v to $\mathrm{e}_k(\theta_\infty)$ in $\overrightarrow{Q}_\infty^{(k)}$. The condition $\mathfrak{s}_{s(r)+1}(\theta_\infty^{(k)}) \prec \mathrm{e}_0(\theta_\infty)$ now has two consequences, as noted in the caption of Figure 4.7:

— First, $\mathrm{e}_k(\theta_\infty)$ belongs to $L(\theta_\infty) \setminus \mathcal{S}_\infty$. Thus the geodesic γ goes from a point of $R(\theta_\infty) \setminus \mathcal{S}_\infty$ to a point of $L(R_\infty) \setminus \mathcal{S}_\infty$, so there exists a vertex v' of γ which belongs to the spine $\mathcal{S}_\infty$ (see the remark we made at the beginning of the proof).

— Second, the set

$$R_\infty^{(k)}(s(r)) \cup \bigcup_{i=0}^{s(r)} L_i(\theta_\infty^{(k)})$$

does not intersect $\mathcal{S}_\infty$, so inclusion (4.19) implies that

$$V\left(B_{\overrightarrow{Q}_\infty^{(k)}}(r)\right) \cap V\left(\mathcal{S}_\infty\right) = \emptyset. \tag{4.20}$$

Putting these two facts together, we get that

$$d_{\overrightarrow{Q}_\infty^{(k)}}(\mathrm{v}, \mathrm{e}_k(\theta_\infty)) \geq d_{\overrightarrow{Q}_\infty^{(k)}}(\mathrm{v}', \mathrm{e}_k(\theta_\infty)) \geq r+1.$$

Similarly, in the second case, conditionally on (4.18), we have

$$V\left(B_{\overleftarrow{Q}_\infty^{(k)}}(r)\right) \cap V\left(R(\theta_\infty)\right) \subset V\left(L_\infty^{(k)}(s(r)) \cup \bigcup_{i=0}^{s(r)} R_i(\theta_\infty^{(k)})\right),$$

and conditionally on $\mathfrak{s}_{s(r)+1}(\theta_\infty^{(-k+1)}) \prec \mathrm{e}_0(\theta_\infty)$, the latter set does not intersect $\mathcal{S}_\infty$, so equation (4.20) still holds. Thus we only have to show that the vertices of $L(\theta_\infty) \setminus \mathcal{S}_\infty$ are at distance at least $r+1$ from e'_k in $\overleftarrow{Q}_\infty^{(k)}$. As above, for every such vertex v, any geodesic path from v to e'_k in $\overleftarrow{Q}_\infty^{(k)}$ intersects $\mathcal{S}_\infty$, hence

$$d_{\overleftarrow{Q}_\infty^{(k)}}(\mathrm{v}, \mathrm{e}_k(\theta_\infty)) \geq r+1.$$

$\square$

From now on, we fix $r \in \mathbb{N}$. The goal of the next sections is to show that the above conditions hold with arbitrarily high probability, for k large enough. For condition (4.17), the main ingredients are the following lemmas:

Lemma 4.4. *Let $s \in \mathbb{N}$ and $\varepsilon > 0$. There exists $s_L = s_L(r, s, \varepsilon)$ such that for all $s' \geq s_L$, there exists $h_L(s', \varepsilon)$ such that for all k large enough, possibly infinite, we have*

$$\mathbb{P}\left(\theta_\infty^{(k)} \notin \mathbb{A}_L(r, s, s', h_L(s', \varepsilon))\right) \leq \varepsilon.$$

Lemma 4.5. *For all $\varepsilon > 0$, there exists $s_R = s_R(r, \varepsilon)$ such that for all k large enough, possibly infinite, we have*

$$\mathbb{P}\left(\overline{\mathcal{A}_{R+}^{(k)}(r, s_R)}\right) \leq \varepsilon,$$

where $\overline{\mathcal{A}_{R+}^{(k)}(r, s_R)}$ denotes the contrary of the event $\mathcal{A}_{R+}^{(k)}(r, s_R)$.

The proofs of these results are given in Sections 4.3 and 4.4, respectively. A first step consists in studying the spine labels of $\overrightarrow{\theta_\infty}$: this is what we do in Section 4.2. In Section 4.5, we finally put all these ingredients together to complete the proof of Proposition 4.1.

4.2 Two properties of the spine labels

In this section, we show two lemmas on the spine labels $S_i(\overrightarrow{\theta_\infty})$. The first one gives an upper bound which holds almost surely, for all i large enough. The second one gives a lower bound which holds with high probability.

Lemma 4.6. *There exists a constant K such that almost surely, for all i large enough, we have*

$$S_i(\overrightarrow{\theta_\infty}) \leq K\sqrt{i \ln(i)}.$$

Proof. Recall that the distribution of $(S_i(\overrightarrow{\theta_\infty}))_{i \geq 0}$ is given in Theorem 1.4. Let $K > 0$ and $i \geq 1$. Recall that

$$\mathbb{P}\left(S_i(\overrightarrow{\theta_\infty}) > K\sqrt{i \ln(i)}\right) = \mathbb{E}\left[M_i \mathbf{1}_{\{\hat{X}_i > K\sqrt{i\ln(i)}\}}\right],$$

where $\hat{X}$ is a random walk with uniform steps in $\{-1, 0, 1\}$ and M is the martingale defined by

$$M_i = \frac{f(\hat{X}_i)}{f(\hat{X}_0)} \prod_{j=0}^{i-1} w(\hat{X}_j).$$

Note that $M_i \leq f(\hat{X}_i)$ almost surely. Thus, for all $\lambda > 0$, we have

$$\begin{aligned}\mathbb{P}\left(S_i(\overrightarrow{\theta_\infty}) > K\sqrt{i \ln(i)}\right) &\leq \mathbb{E}\left[M_i e^{-\lambda K\sqrt{i\ln(i)}} e^{\lambda \hat{X}_i}\right] \\ &\leq C e^{-\lambda K \sqrt{i \ln(i)}} \mathbb{E}\left[\hat{X}_i^4 e^{\lambda \hat{X}_i}\right], \qquad (4.21)\end{aligned}$$

where C denotes a constant such that $f(x) = x(x+3)(2x+3) \leq Cx^4$ for every $x \geq 0$. Now $\mathbb{E}\left[\hat{X}_i^4 e^{\lambda \hat{X}_i}\right]$ is the fourth derivative of $\mathbb{E}\left[e^{\lambda \hat{X}_i}\right]$, and we have

$$\mathbb{E}\left[e^{\lambda \hat{X}_i}\right] = e^{i\psi(\lambda)},$$

where $\psi(\lambda)$ denotes the Laplace transform of a uniform random variable in $\{-1, 0, 1\}$. Now we have

$$\psi(\lambda) = \ln\left(\frac{1 + 2\cosh(\lambda)}{3}\right) \leq c\lambda^2$$

for a suitable constant $c > 0$, and that the first four derivatives of ψ are bounded. Therefore, there exists a positive constant C' such that

$$C\mathbb{E}\left[\hat{X}_i^4 e^{\lambda \hat{X}_i}\right] \le C' i^4 e^{i\psi(\lambda)} \le C' i^4 e^{ci\lambda^2} \qquad \forall \lambda > 0.$$

Putting this together with (4.21), we get

$$\mathbb{P}\left(S_i(\overrightarrow{\theta_\infty}) > K\sqrt{i\ln(i)}\right) \le C' i^4 e^{ci\lambda^2 - \lambda K\sqrt{i\ln(i)}} \qquad \forall \lambda > 0.$$

Choosing the optimal value $\lambda = K\sqrt{\ln(i)}/(2c\sqrt{i})$ gives

$$\mathbb{P}\left(S_i(\overrightarrow{\theta_\infty}) > K\sqrt{i\ln(i)}\right) \le C' i^4 e^{-(K^2/2c)\ln(i)} = C' i^{4-K^2/2c}.$$

As a consequence, for all K large enough (such that $4-K^2/2c < -1$), the sum of the probabilities $\mathbb{P}\left(S_i(\overrightarrow{\theta_\infty}) > K\sqrt{i\ln(i)}\right)$ is finite. Applying the Borel–Cantelli lemma concludes the proof. $\square$

Lemma 4.7. *For all $\eta > 0$, there exists $\delta > 0$ such that for all s large enough, we have*

$$\mathbb{P}\left(\exists i \ge \lfloor \eta s \rfloor : S_i(\overrightarrow{\theta_\infty}) < \lfloor \delta\sqrt{s} \rfloor\right) \le \eta.$$

Proof. Since $(S_i(\overrightarrow{\theta_\infty}))_{i\ge 0}$ has the same distribution as $(\tilde{X}_i)_{i\ge 0}$ with $\tilde{X}_0 = 1$, it is enough to show that

$$\mathbb{P}_1\left(\exists i \ge \lfloor \eta s \rfloor : \tilde{X}_i < \lfloor \delta\sqrt{s} \rfloor\right) \le \eta.$$

Recall that, as stated in the introduction, we have the convergence

$$\left(\frac{1}{\sqrt{n}}\tilde{X}_{\lfloor nt \rfloor}\right)_{t\ge 0} \xrightarrow[n\to\infty]{(d)} (Z_{2t/3})_{t\ge 0},$$

where Z denotes a seven-dimensional Bessel process. As a consequence, there exists constants $\delta_1 > 0$ and $s_1 \in \mathbb{N}$ such that, for all $s \ge s_1$, we have

$$\mathbb{P}_1\left(\tilde{X}_{\lfloor \eta s \rfloor} \le \sqrt{\eta s}\cdot\delta_1\right) \le \frac{\eta}{2}.$$

Fix $s \ge s_1$. Using the Markov property at time $\lfloor \eta s \rfloor$, for any $\delta > 0$, we can now write

$$\begin{aligned}
\mathbb{P}_1\left(\exists i \ge \lfloor \eta s \rfloor : \tilde{X}_i < \lfloor \delta\sqrt{s} \rfloor\right) &\le \frac{\eta}{2} + \sum_{x \ge \sqrt{\eta s}\cdot\delta_1} \mathbb{P}_1(\tilde{X}_{\lfloor \eta s \rfloor} = x)\mathbb{P}_x\left(\exists i \ge 0 : \tilde{X}_i < \lfloor \delta\sqrt{s} \rfloor\right) \\
&\le \frac{\eta}{2} + \sum_{x \ge \sqrt{\eta s}\cdot\delta_1} \mathbb{P}_1(\tilde{X}_{\lfloor \eta s \rfloor} = x)\mathbb{P}_x\left(\exists i \ge 0 : \tilde{X}_i < \frac{\delta}{\delta_1\sqrt{\eta}}x\right) \\
&= \frac{\eta}{2} + \sum_{x \ge \sqrt{\eta s}\cdot\delta_1} \mathbb{P}_1(\tilde{X}_{\lfloor \eta s \rfloor} = x)\mathbb{P}_x\left(\tilde{T}_{\lfloor \delta x/\delta_1\sqrt{\eta} \rfloor} < \infty\right), \qquad (4.22)
\end{aligned}$$

where $\tilde{T}_{x'}$ denotes the first hitting time of x' for $\tilde{X}$. It was shown in the proof of Lemma 2.2 that for all $x \ge x'$, we have

$$\mathbb{P}_x(\tilde{T}_{x'} < \infty) = \frac{h(x')}{h(x)},$$

for a given non-constant polynomial h. Thus there exists constants δ_2 and x_2 such that for all $x \ge x_2$, we have

$$\mathbb{P}_x(\tilde{T}_{\lfloor \delta_2 x \rfloor} < \infty) \le \frac{\eta}{2}.$$

Putting this together with (4.22), for $\delta = \delta_2\sqrt{\delta_1\eta}$ and $s \ge s_1 \wedge (x_2^2/(\eta\delta_1^2))$, we get

$$\mathbb{P}_1\left(\exists i \ge \lfloor \eta s \rfloor : \tilde{X}_i < \lfloor \delta\sqrt{s} \rfloor\right) \le \frac{\eta}{2}\left(1 + \sum_{x \ge \sqrt{\eta s}\cdot\delta_1} \mathbb{P}_1(\tilde{X}_{\lfloor \eta s \rfloor} = x)\right) \le \eta.$$

$\square$

4.3 Proof of the left-hand condition

In this section, we give the proof of Lemma 4.4. This result mainly uses the upper bound on the spine labels of $\overrightarrow{\theta_\infty}$, and the explicit expressions of the distribution of $L_i(\overrightarrow{\theta_\infty})$, for $i \geq 0$.

Proof of Lemma 4.4. Since for all s, s', h, $\mathbb{S} \setminus \mathbb{A}_L(r,s,s',h)$ is a closed set, we have

$$\limsup_{k\to\infty} \mathbb{P}\left(\theta_\infty^{(k)} \notin \mathbb{A}_L(r,s,s',h)\right) \leq \mathbb{P}\left(\overrightarrow{\theta_\infty} \notin \mathbb{A}_L(r,s,s',h)\right),$$

so it is enough to show that the Lemma holds with $\overrightarrow{\theta_\infty}$ instead of $\theta_\infty^{(k)}$. For all $s, s', h \in \mathbb{N}$, we have

$$\mathbb{P}\left(\overrightarrow{\theta_\infty} \notin \mathbb{A}_L(r,s,s',h)\right) \leq p_{r,s,s'} + \mathbb{P}\left(\exists i \leq s' : L_i(T) \text{ has height greater than } h - s'\right),$$

where

$$p_{r,s,s'} := \mathbb{P}\left(\forall i \in \left\{s+1,\dots,s'\right\},\ \min_{\mathrm{x}\in L_i(\overrightarrow{\theta_\infty})} \overrightarrow{l_\infty}(\mathrm{x}) > -r\right).$$

Since for all s', there exists h such that the maximum of the heights of the Galton–Watson trees $L_0(\overrightarrow{\theta_\infty}), \dots L_{s'}(\overrightarrow{\theta_\infty})$ is less than $h - s'$ with probability greater than $1-\varepsilon/2$, it is enough to prove that the probabilities $p_{r,s,s'}$ converge to 0 as $s' \to \infty$.

We first rewrite $p_{r,s,s'}$ using the spine-labels $S_i(\overrightarrow{\theta_\infty})$:

$$\begin{aligned} p_{r,s,s'} &= \mathbb{E}\left[\prod_{i=s+1}^{s'} \rho_{(S_i(\overrightarrow{\theta_\infty}))}\left\{(T,l) \in \mathbb{T} : l > -r\right\}\right] \\ &= \mathbb{E}\left[\prod_{i=s+1}^{s'} \rho_{(r+S_i(\overrightarrow{\theta_\infty}))}\left(\mathbb{T}^+\right)\right] = \mathbb{E}\left[\prod_{i=s+1}^{s'} w(r + S_i(\overrightarrow{\theta_\infty}))\right]. \end{aligned}$$

The above product is almost surely decreasing as $s' \to \infty$. Therefore, we only have to show that

$$\prod_{i=s+1}^{s'} w(r + S_i(\overrightarrow{\theta_\infty})) \xrightarrow[s'\to\infty]{} 0,$$

or equivalently that

$$\sum_{i=s+1}^{s'} -\ln\left(w(r + S_i(\overrightarrow{\theta_\infty}))\right) \xrightarrow[s'\to\infty]{} +\infty \tag{4.23}$$

almost surely. Since $S_i(\overrightarrow{\theta_\infty}) \to +\infty$ almost surely, we can use the estimate

$$w(x) = 1 - \frac{2}{x^2} + o\left(\frac{1}{x^2}\right).$$

This yields

$$-\ln\left(w(r + S_i(\overrightarrow{\theta_\infty}))\right) \sim_{i\to\infty} \frac{2}{\left(r + S_i(\overrightarrow{\theta_\infty})\right)^2}.$$

Lemma 4.6 now ensures that the right-hand term is a.s. larger than $2/(K^2 i \ln(i))$ for all i large enough, hence the a.s. divergence (4.23). $\square$

4.4 Proof of the right-hand condition

This section is devoted to the proof of Lemma 4.5. Note that the structure of the proof is close to Ménard [52]. More precisely, the lower bound we already proved in Lemma 4.7 corresponds to a result Ménard obtains by putting together Lemma 2 and Proposition 5 of [52], and Lemma 4.10 corresponds to Lemma 5 of [52].

We begin by computing the probability $\mathbb{P}(R_\infty^{(k)}(s) = \theta^*)$, and some conditional probabilities on this event, for all suitable trees θ^*. More precisely, let $\mathbb{T}_{[s]}^R$ denote the set of the labeled trees $(T, l) \in \mathbb{T}(0)$ such that:

— The root of T has exactly one offspring.

— All labels in T are positive, except the root-label.

— The height of T is s.

— There are no vertices on the left of the path from the root to x_s, where x_s denotes the leftmost vertex having height s. In other words, if $\mathrm{x}_0, \ldots, \mathrm{x}_s$ are the vertices on the path from the root to x_s, then for all $\mathrm{x} \in T^* \setminus \{\mathrm{x}_0, \ldots, \mathrm{x}_s\}$, we have $\mathrm{x} > \mathrm{x}_s$ (where $<$ denotes the depth-first order).

Fix $\theta^* \in \mathbb{T}_{[s]}^R$. We let $\mathrm{x}_s = \mathrm{y}_1 < \ldots < \mathrm{y}_{n^*}$ denote the vertices of θ^* which have height s. For all $i \in \{0, \ldots, s-1\}$, we let τ_i^* denote the subtree formed by the vertices $\mathrm{x} \in T^*$ such that $\mathrm{x}_i \preceq \mathrm{x}$ and $\mathrm{x}_{i+1} \not\preceq \mathrm{x}$ (note that $\tau_0^* = \{\mathrm{x}_0\}$). Finally, for all suitable i, we let $x_i = l^*(\mathrm{x}_i)$ and $y_i = l^*(\mathrm{y}_i)$. We have the following results:

Lemma 4.8. *Let $k > s + r$. With the above notation, we have*

$$\mathbb{P}\left(R_\infty^{(k)}(s) = \theta^*\right) = W_s(y_1, k) \frac{2^{n^*-1}}{6^{s-1} 12^{|T^*|-s}} \prod_{j=2}^{n^*} w(y_j), \tag{4.24}$$

where

$$W_s(x, k) = \frac{f(x)}{f(1)} \left(1 - \frac{C_x - s + 1}{10(k+1)(k+2)}\right) \qquad \forall x \leq s < k.$$

Moreover, this yields the conditional probabilities

$$\mathbb{P}\left(\min_{\bigcup_{i \geq s} \tau_{\infty,i}^{(k)}} l_\infty^{(k)} > r \middle| R_\infty^{(k)}(s) = \theta^*\right) = \frac{W_s(y_1 - r, k - r)}{W_s(y_1, k)} \tag{4.25}$$

and

$$\mathbb{P}\left(\min_{\mathrm{y}_j \preceq \mathrm{v}} l_\infty^{(k)}(\mathrm{v}) > r \middle| R_\infty^{(k)}(s) = \theta^*\right) = \frac{w(y_j - r)}{w(y_j)} \qquad \forall j \in \{2, \ldots, n^*\}. \tag{4.26}$$

Note that it is easy to see that these equations also hold for $k = \infty$, with $RB_{\overrightarrow{\theta_\infty}}(s)$ instead of $R_\infty^{(k)}(s)$ and $W_s(x, \infty) = f(x)/f(1)$ for all $x \leq s$.

Proof. Note that we have $x_s = y_1$; in the first two steps of the proof, it is more natural to use the notation x_s. The characterization of the distribution of $\theta_\infty^{(k)}$ given in Proposition 2.1 yields

$$\mathbb{P}\left(R_\infty^{(k)}(s) = \theta^*\right) = \mathbb{P}\left((X_{\infty,1}^{(k)}, \ldots, X_{\infty,s}^{(k)}) = (x_1, \ldots, x_s)\right) \prod_{i=1}^{s-1} \rho_{(x_i)}^+ \left(\theta : B_\theta(s-i) = \tau_i^*\right). \tag{4.27}$$

Furthermore, the computations of Section 2.1 show that

$$\begin{aligned}\mathbb{P}\left(X_{\infty,i}^{(k)}=x_i,\ \forall i\leq s\right)&=\sum_{m\geq s}\frac{m+1}{3^s}\prod_{i=1}^{s-1}w(x_i)\ \mathbb{E}_{x_s}\left[\prod_{i=0}^{m-s}w(\hat{X}_i)\mathbf{1}_{\{\hat{X}_{m-s}=k\}}\right]\\&=3^{-s}\left(\prod_{i=1}^{s-1}w(x_i)\right)\frac{w(k)f(x_s)}{f(k)}\sum_{m\geq 0}(m+s+1)\mathbb{P}_{x_s}\left(\tilde{X}_m=k\right)\\&=3^{-s}\left(\prod_{i=1}^{s-1}w(x_i)\right)\frac{w(k)f(x_s)}{f(k)}\left(H_{x_s}^*(k)+(s-1)H_{x_s}(k)\right).\end{aligned}$$

Using the expressions obtained in Lemma 2.2 and Lemma 2.3, and the hypothesis $s<k$, this gives

$$\mathbb{P}\left(X_{\infty,i}^{(k)}=x_i,\ \forall i\leq s\right)=3^{-s+1}\left(\prod_{i=1}^{s-1}w(x_i)\right)\frac{f(x_s)}{f(1)}\left(1-\frac{C_{x_s}-s+1}{10(k+1)(k+2)}\right).$$

Besides, for all $i\leq s-1$, we have

$$\rho_{(x_i)}^{+}\left(\theta:B_\theta(s-i)=\tau_i^*\right)=\frac{1}{2w(x_i)12^{|\tau_i^*|}}\prod_{j:\mathrm{v}_j\in\tau_i^*}2w(y_j).$$

Equation (4.27) can now be rewritten as

$$\begin{aligned}\mathbb{P}\left(R_\infty^{(k)}(s)=\theta^*\right)&=\frac{f(x_s)}{f(1)}\left(1-\frac{C_{x_s}-s+1}{10(k+1)(k+2)}\right)\prod_{i=1}^{s-1}\left(\frac{w(x_i)}{6w(x_i)12^{|\tau_i^*|}}\prod_{j:\mathrm{v}_j\in\tau_i^*}2w(y_j)\right)\\&=\frac{f(x_s)}{f(1)}\left(1-\frac{C_{x_s}-s+1}{10(k+1)(k+2)}\right)\frac{2^{n^*-1}}{6^{s-1}12^{|T^*|-s}}\prod_{j=2}^{n^*}w(y_j),\end{aligned}$$

hence the first result of the lemma.

To get the conditional probability (4.25), we have to compute

$$\mathbb{P}\left(\left(R_\infty^{(k)}(s)=\theta^*\right)\cap\left(\min_{\bigcup_{i\geq s}\tau_{\infty,i}^{(k)}}l_\infty^{(k)}>r\right)\right).$$

Using the same decomposition as above, this probability can be written as

$$\sum_{m\geq s}\frac{m+1}{3^m}\left(\prod_{i=1}^{s-1}w(x_i)\right)\sum_{\underline{x}'\in\mathcal{M}_{m-s,x_s\to k}^+}\left(\prod_{i=0}^{m-s}\rho_{(x_i')}^{+}\left((T,l):\min_T l>r\right)\right)$$
$$\prod_{i=1}^{s-1}\rho_{(x_i)}^{+}\left(\theta:B_\theta(s-i)=\tau_i^*\right),$$

or equivalently,

$$\left(\sum_{m\geq 0}\frac{m+s+1}{3^m}\sum_{\underline{x}'\in\mathcal{M}_{m,x_s-r\to k-r}^+}\prod_{i=0}^{m}w(x_i')\right)\frac{2^{n^*-1}}{6^{s-1}12^{|T^*|-s}}\prod_{j=2}^{n^*}w(y_j).$$

Thus, we get

$$\mathbb{P}\left(\left(R_\infty^{(k)}(s)=\theta^*\right)\cap\left(\min_{\bigcup_{i\geq s}\tau_{\infty,i}^{(k)}} l_\infty^{(k)}>r\right)\right)$$
$$=\frac{w(k-r)f(x_s-r)}{3f(k-r)}\left(H^*_{x_s-r}(k-r)+(s-1)H_{x_s-r}(k-r)\right)\frac{2^{n^*-1}}{6^{s-1}12^{|T^*|-s}}\prod_{j=2}^{n^*}w(y_j)$$
$$=\frac{f(x_s-r)}{f(1)}\left(1-\frac{C_{x_s-r}-s+1}{10(k-r+1)(k-r+2)}\right)\frac{2^{n^*-1}}{6^{s-1}12^{|T^*|-s}}\prod_{j=2}^{n^*}w(y_j).$$

This completes the proof of equation (4.25).

Finally, for all $j^*\in\{2,\ldots,n^*\}$, we have

$$\mathbb{P}\left(\left(R_\infty^{(k)}(s)=\theta^*\right)\cap\left(\min_{\mathrm{y}_j\preceq\mathrm{v}} l_\infty^{(k)}(\mathrm{v})>r\right)\right)$$
$$=3^{-s+1}\left(\prod_{i=1}^{s-1}w(x_i)\right)W(y_1,k)\prod_{i=1}^{s-1}\frac{1}{2w(x_i)12^{|\tau_i^*|}}\left(\prod_{\substack{j:\mathrm{v}_j\in\tau_i^*\\ j\neq j^*}}2w(y_j)\right)\times 2w(y_{j^*}-r)$$
$$=W_s(y_1,k)\frac{2^{n^*-1}}{6^{s-1}12^{|T^*|-s}}w(y_{j^*}-r)\prod_{\substack{2\leq j\leq n^*\\ j\neq j^*}}w(y_j),$$

hence equation (4.26). □

The second step consists in studying the vertices of $R_\infty^{(k)}$ which are exactly at height s: we give an upper bound on the expectation of the number of such vertices, and show that with high probability, for k large enough, these vertices have labels greater than s^α, for $\alpha\in(0,1/2)$. Precise statements are given in Lemmas 4.9 and 4.10 below. Note that for all k, we have

$$\partial R_\infty^{(k)}(s):=\{\mathrm{v}\in R_\infty^{(k)}:\mathrm{v}\text{ has height }s\}\subset\partial RB_{\theta_\infty^{(k)}}(s).$$

Lemma 4.9. *For all $s\geq 1$ and $k\in\mathbb{N}$, we have*

$$\mathbb{E}\left[\#\partial R_\infty^{(k)}(s)\right]\leq s.$$

Proof. For all $s\geq 1$ and $k\in\mathbb{N}$, we have

$$\mathbb{E}\left[\#\partial R_\infty^{(k)}(s)\right]=\sum_{i=1}^{s}\mathbb{E}\left[\#\partial B_{\tau_{\infty,i}^{(k)}}(s-i)\right]$$
$$=\sum_{i=1}^{s}\mathbb{E}\left[\frac{1}{w(X_i)}\mathbb{E}[\#\partial B_\tau(s-i)]\right],$$

where τ denotes a Galton–Watson tree with offspring distribution Geom(1/2). For all $h\geq 0$, we have $\mathbb{E}[\#\partial B_\tau(h)]=1$. As a consequence, the above equality gives

$$\mathbb{E}\left[\#\partial R_\infty^{(k)}(s)\right]=\sum_{i=1}^{s}\mathbb{E}\left[\frac{1}{w(X_i)}\right]\leq\sum_{i=1}^{s}1=s.$$

□

We now consider the set

$$\tilde{\mathbb{A}}_{R+}(r,s,\alpha)=\left\{(T,l)\in\mathbb{S}:\forall\mathrm{v}\in\partial RB_T(s),\ l(\mathrm{v})>\lfloor s^\alpha\rfloor\right\}.$$

Lemma 4.10. *Fix $\alpha < 1/2$. For all s large enough, there exists $k_1(s)$ such that for all $k \geq k_1(s)$, possibly infinite, we have*

$$\mathbb{P}\left(\theta_\infty^{(k)} \notin \tilde{\mathbb{A}}_{R+}(r,s,\alpha)\right) \leq \varepsilon.$$

Proof. First note that since $\mathbb{S} \setminus \tilde{\mathbb{A}}_{R+}(r,s,\alpha)$ is a closed set, we have

$$\limsup_{k\to\infty} \mathbb{P}\left(\theta_\infty^{(k)} \notin \tilde{\mathbb{A}}_{R+}(r,s,\alpha)\right) \leq \mathbb{P}\left(\overrightarrow{\theta_\infty} \notin \tilde{\mathbb{A}}_{R+}(r,s,\alpha)\right),$$

so it is enough to show that the property holds for $k=\infty$. Moreover, the same arguments as in the proof of [52, Lemma 5] show that for all $\eta \in (0,1/2)$, for all s large enough, we have

$$\mathbb{P}\left(\exists i \leq \lfloor \eta s \rfloor - 1 : R_i(\overrightarrow{\theta_\infty}) \cap \partial B_{\overrightarrow{\theta_\infty}}(s) \neq \emptyset\right) \leq 4\eta.$$

Thus, letting $I_\eta(s) = \{\lfloor \eta s \rfloor, \dots, s\}$, we have

$$\mathbb{P}\left(\overrightarrow{\theta_\infty} \notin \tilde{\mathbb{A}}_{R+}(r,s,\alpha)\right) \leq 4\eta + \mathbb{P}\left(\exists i \in I_\eta(s) : \min_{R_i(\overrightarrow{\theta_\infty})} \overrightarrow{l_\infty} \leq \lfloor s^\alpha \rfloor\right).$$

Lemma 4.7 now ensures that for $\delta > 0$ and s large enough, this probability is less than

$$5\eta + \mathbb{P}\left(\left(\exists i \in I_\eta(s) : \min_{R_i(\overrightarrow{\theta_\infty})} \overrightarrow{l_\infty} \leq \lfloor s^\alpha \rfloor\right) \cap \left(\forall i \in I_\eta(s), S_i(\overrightarrow{\theta_\infty}) \geq \lfloor \delta\sqrt{s} \rfloor\right)\right). \tag{4.28}$$

For all $(x_i)_{\lfloor \eta s \rfloor \leq i \leq s}$, we have

$$\begin{aligned}
&\mathbb{P}\left(\exists i \in I_\eta(s) : \min_{R_i(\overrightarrow{\theta_\infty})} \overrightarrow{l_\infty} \leq \lfloor s^\alpha \rfloor \,\middle|\, S_i(\overrightarrow{\theta_\infty}) = x_i \ \forall i \in I_\eta(s)\right) \\
&\qquad\qquad \leq \sum_{i=\lfloor \eta s \rfloor}^{s} \mathbb{P}\left(\min_{R_i(\overrightarrow{\theta_\infty})} \overrightarrow{l_\infty} \leq \lfloor s^\alpha \rfloor \,\middle|\, S_i(\overrightarrow{\theta_\infty}) = x_i\right) \\
&\qquad\qquad \leq \sum_{i=\lfloor \eta s \rfloor}^{s} \rho^+_{(x_i)}\left((T,l) \in \mathbb{T}^+ : \min_T l \leq \lfloor s^\alpha \rfloor\right) \\
&\qquad\qquad \leq \sum_{i=\lfloor \eta s \rfloor}^{s} \frac{w(x_i) - w(x_i - \lfloor s^\alpha \rfloor)}{w(x_i)}.
\end{aligned}$$

Furthermore, if we choose the integers x_i in such a way that $x_i \geq \lfloor \delta\sqrt{s} \rfloor$ for all i, we have

$$\frac{w(x_i) - w(x_i - \lfloor s^\alpha \rfloor)}{w(x_i)} = \frac{4\lfloor s^\alpha \rfloor}{x_i^3} + o\left(\frac{s^\alpha}{x_i^3}\right) \leq \frac{4s^{\alpha-3/2}}{\delta^3} + o\left(s^{\alpha-3/2}\right),$$

so

$$\mathbb{P}\left(\exists i \in I_\eta(s) : \min_{R_i(\overrightarrow{\theta_\infty})} \overrightarrow{l_\infty} \leq \lfloor s^\alpha \rfloor \,\middle|\, S_i(\overrightarrow{\theta_\infty}) = x_i \ \forall i \in I_\eta(s)\right) \leq \frac{4s^{\alpha-1/2}}{\delta^3} + o\left(s^{\alpha-1/2}\right).$$

Putting this together with (4.28), we finally get

$$\mathbb{P}\left(\overrightarrow{\theta_\infty} \notin \tilde{\mathbb{A}}_{R+}(r,s,\alpha)\right) \leq 5\eta + \frac{4s^{\alpha-1/2}}{\delta^3} + o\left(s^{\alpha-1/2}\right) \leq 6\eta$$

for all s large enough. $\square$

We are now ready to give the proof of Lemma 4.5.

Proof of Lemma 4.5. Fix $\alpha \in (1/3, 1/2)$. Lemma 4.10 show that for all s large enough and $k \geq k_1(s)$ (possibly infinite), we have

$$\mathbb{P}\left(\overline{\mathcal{A}_{R+}^{(k)}(r,s)}\right) \leq 2\varepsilon + \mathbb{P}\left(\overline{\mathcal{A}_{R+}^{(k)}(r,s)} \cap \left(\theta_\infty^{(k)} \in \tilde{\mathbb{A}}_{R+}(r,s,\alpha)\right)\right).$$

Letting

$$\Theta(s,\alpha) = \left\{(T,l) \in \mathbb{T}_{[s]}^R : \min_{\partial B_T(s)} l > \lfloor s^\alpha \rfloor\right\},$$

we get that

$$\mathbb{P}\left(\overline{\mathcal{A}_{R+}^{(k)}(r,s)}\right) \leq 2\varepsilon + \sum_{\theta^* \in \Theta(s,\alpha)} \mathbb{P}\left(\left(R_\infty^{(k)}(s) = \theta^*\right) \cap \overline{\mathcal{A}_{R+}^{(k)}(r,s)}\right). \tag{4.29}$$

Fix $\theta^* \in \Theta_\varepsilon(s,\alpha)$, and let $y_1, \ldots, y_{n^*}$ denote the labels of the vertices of height s (from left to right) in θ^*. Note that the condition $\theta^* \in \Theta_\varepsilon(s,\alpha)$ means that we have $\lfloor s^\alpha \rfloor < y_i \leq s$ for all $i \in \{1, \ldots, n^*\}$. Moreover, equations (4.25) and (4.26) show that

$$\mathbb{P}\left(\overline{\mathcal{A}_{R+}^{(k)}(r,s)} \,\middle|\, R_\infty^{(k)}(s) = \theta^*\right) \leq 1 - \frac{W_s(y_1 - r, k - r)}{W_s(y_1, k)} + \sum_{j=2}^{n^*}\left(1 - \frac{w(y_j - r)}{w(y_j)}\right).$$

For all $y \leq s$, we have

$$\frac{W_s(y-r,k-r)}{W_s(y,k)} = \frac{f(y-r)}{f(y)}\left(1 + \frac{\frac{C_{y-r}-s+1}{10(k-r+1)(k-r+2)} - \frac{C_y - s+1}{10(k+1)(k+2)}}{1 - \frac{C_y-s+1}{10(k+1)(k+2)}}\right) \geq \frac{f(y-r)}{f(y)},$$

so

$$0 \leq 1 - \frac{W_s(y-r,k-r)}{W_s(y,k)} \leq 1 - \frac{f(y-r)}{f(y)} \leq \varepsilon$$

for all s large enough and $y \in \{\lfloor s^\alpha \rfloor, \ldots, s\}$. Besides, uniformly in $y > \lfloor s^\alpha \rfloor$, we have

$$1 - \frac{w(y-r)}{w(y)} \leq \frac{4r}{s^{3\alpha}} + o\left(\frac{r}{s^{3\alpha}}\right).$$

This yields

$$\mathbb{P}\left(\overline{\mathcal{A}_{R+}^{(k)}(r,s)} \,\middle|\, R_\infty^{(k)}(s) = \theta^*\right) \leq \varepsilon + n^*\left(\frac{4r}{s^{3\alpha}} + o\left(\frac{r}{s^{3\alpha}}\right)\right).$$

Putting this into (4.29), we obtain

$$\begin{aligned}\mathbb{P}\left(\overline{\mathcal{A}_{R+}^{(k)}(r,s)}\right) &\leq 3\varepsilon + \left(\frac{4r}{s^{3\alpha}} + o\left(\frac{r}{s^{3\alpha}}\right)\right)\sum_{\theta^* \in \Theta(s,\alpha)} \#\partial B_{\theta^*}(s)\mathbb{P}\left(R_\infty^{(k)}(s) = \theta^*\right) \\ &= 3\varepsilon + \left(\frac{4r}{s^{3\alpha}} + o\left(\frac{r}{s^{3\alpha}}\right)\right)\mathbb{E}\left[\#\partial R_\infty^{(k)}(s)\right].\end{aligned}$$

Lemma 4.9 now implies that

$$\mathbb{P}\left(\overline{\mathcal{A}_{R+}^{(k)}(r,s)}\right) \leq 3\varepsilon + \frac{4r}{s^{3\alpha-1}} + o\left(\frac{r}{s^{3\alpha-1}}\right).$$

Since we took $\alpha > 1/3$, this concludes the proof. $\square$

4.5 Proof of Proposition 4.1

We can now prove Proposition 4.1 by putting together the results of Lemmas 4.4, 4.5 and 4.3, and using the symmetry between the definitions of $\theta_\infty^{(k)}$ and $\theta_\infty^{(-k)}$.

Proof of Proposition 4.1. Let $r \in \mathbb{N}$. For all $\varepsilon \geq 0$, we consider the sequences $(s_\varepsilon(r'))_{r'\geq 0}$ and $(h_\varepsilon(r'))_{r'\geq 1}$ defined by $s_\varepsilon(0) = 0$, and for all $r' \geq 1$:

$$s_\varepsilon(r') = s_R(r', 2^{-r'-1}\varepsilon) \vee s_L(r', s_\varepsilon(r'-1), 2^{-r'-1}\varepsilon)$$
$$h_\varepsilon(r') = h_L(s_\varepsilon(r'), 2^{-r'-1}\varepsilon),$$

where s_L, s_R and h_R are the quantities introduced in Lemmas 4.4 and 4.5. Note that for all r', we have $\mathcal{A}_{R+}(r', s_\varepsilon(r')) \subset \mathcal{A}_{R+}(r', s_R(r', 2^{-r'-1}\varepsilon))$. Thus, Lemmas 4.4 and 4.5 show that for all $r' \in \mathbb{N}$, for all k large enough, we have

$$\mathbb{P}\left(\left(\theta_\infty^{(k)} \notin \mathbb{A}_L(r', s_\varepsilon(r'-1), s_\varepsilon(r'), h_\varepsilon(r'))\right) \cup \overline{\mathcal{A}_{R+}^{(k)}(r', s_\varepsilon(r'))}\right) \leq 2^{-r'}\varepsilon,$$

and as a consequence,

$$\mathbb{P}\left(\bigcup_{r'=1}^{r}\left(\theta_\infty^{(k)} \notin \mathbb{A}_L(r', s_\varepsilon(r'-1), s_\varepsilon(r'), h_\varepsilon(r'))\right) \cup \overline{\mathcal{A}_{R+}^{(k)}(r', s_\varepsilon(r'))}\right) \leq \varepsilon$$

for all k large enough. Moreover, recalling the notation of Proposition 2.1, we have

$$\mathfrak{s}_{s_\varepsilon(r)+1}(\theta_\infty^{(k)}) \not\prec \mathrm{e}_0(\theta_\infty)$$

if and only if $I_\infty^{(k)} \leq s_\varepsilon(r)$, which happens with probability at most $s_\varepsilon(r)/(k+1)$. Therefore, for all k large enough, the conditions stated in the first part of Lemma 4.3 hold with probability at least $1 - 2\varepsilon$.

Finally, we can see from the symmetry between the definitions of $\theta_\infty^{(k)}$ and $\theta_\infty^{(-k)}$ that for all $r', s, s', h \in \mathbb{N}$, we have

$$\mathbb{P}\left(\theta_\infty^{(-k+1)} \notin \mathbb{A}_R(r', s, s', h)\right) = \mathbb{P}\left(\theta_\infty^{(k-1)} \notin \mathbb{A}_L(r'+1, s, s', h)\right)$$

and

$$\mathbb{P}\left(\overline{\mathcal{A}_{L+}^{(-k+1)}(r', s)}\right) = \mathbb{P}\left(\overline{\mathcal{A}_{R+}^{(k-1)}(r'-1, s)}\right).$$

Thus, letting $\tilde{s}_\varepsilon(0) = 0$ and, for all $r' \geq 1$,

$$\tilde{s}_\varepsilon(r') = s_R(r'-1, 2^{-r'-1}\varepsilon) \vee s_L(r'+1, \tilde{s}_\varepsilon(r'-1), 2^{-r'-1}\varepsilon)$$
$$\tilde{h}_\varepsilon(r') = h_L(\tilde{s}_\varepsilon(r'), 2^{-r'-1}\varepsilon),$$

the probability

$$\mathbb{P}\left(\bigcap_{r'=1}^{r}\left(\theta_\infty^{(-k+1)} \notin \mathbb{A}_R(r', \tilde{s}_\varepsilon(r'-1), \tilde{s}_\varepsilon(r'), \tilde{h}_\varepsilon(r'))\right) \cap \mathcal{A}_{L+}^{(-k+1)}(r', \tilde{s}_\varepsilon(r'))\right)$$

is equal to

$$\mathbb{P}\left(\bigcap_{r'=1}^{r}\left(\theta_\infty^{(k-1)} \notin \mathbb{A}_L(r'+1, \tilde{s}_\varepsilon(r'-1), \tilde{s}_\varepsilon(r'), \tilde{h}_\varepsilon(r'))\right) \cap \mathcal{A}_{R+}^{(k-1)}(r'-1, \tilde{s}_\varepsilon(r'))\right) \leq \varepsilon.$$

Similarly as above, this implies that the conditions stated in the second part of Lemma 4.3 hold with probability at least $1 - 2\varepsilon$.

Therefore, Lemma 4.3 shows that for $h = h_\varepsilon(r) \vee \tilde{h}_\varepsilon(r)$, the inclusions (4.14) and (4.15) hold with probability at least $1 - 4\varepsilon$, for all k large enough. $\square$

Bibliography

[1] Romain Abraham and Jean François Delmas. Beta-coalescents and stable Galton-Watson trees. 2013.

[2] Romain Abraham and Jean François Delmas. The forest associated with the record process on a Lévy tree. *Stochastic Processes and their Applications*, (123):3497–3517, 2013.

[3] Romain Abraham, Jean François Delmas, and Guillaume Voisin. Pruning a Lévy continuum random tree. *Electronic Journal of Probability*, 15:1429–1473, 2010.

[4] Louigi Addario-Berry, Nicolas Broutin, and Cecilia Holmgren. Cutting down trees with a Markow chainsaw. *The Annals of Applied Probability*, 2012.

[5] David Aldous. The continuum random tree I. *Annals of Probability*, 19(1):1–28, 1991.

[6] David Aldous. The continuum random tree II: an overview. *Stochastic Analysis*, pages 23–70, 1991.

[7] David Aldous. The continuum random tree III. *The Annals of Probability*, 21(1):248–289, 1993.

[8] David Aldous. Recursive self-similarity for random trees, random triangulations and brownian excursion. *The Annals of Probability*, 22(2):527–545, 1994.

[9] David Aldous and Jim Pitman. The standard additive coalescent. *The Annals of Probability*, 26(4):1723–1726, 1998.

[10] Omer Angel. Growth and percolation on the uniform infinite planar triangulation. *Geometric & Functional Analysis*, 13(5):935–974, 2003.

[11] Omer Angel. Scaling of percolation on infinite planar maps. 2005. preprint.

[12] Omer Angel and Oded Schramm. Uniform infinite planar triangulations. *Communications in Mathematical Physics*, 241(2-3):191–213, October 2003.

[13] Jean Bertoin. *Lévy processes.* Cambridge University Press, 1998.

[14] Jean Bertoin. A fragmentation process connected to Brownian motion. *Probability Theory and Related Fields*, 117:289–301, 2000.

[15] Jean Bertoin. Self-similar fragmentations. *Annales de l'institut Henri Poincaré, Probabilités et Statistiques*, 38:319–340, 2002.

[16] Jean Bertoin. *Random fragmentation and coagulation processes.* Cambridge studies in advanced mathematics. Cambridge University Press, 2006.

[17] Jean Bertoin. Fires on trees. *Annales de l'institut Henri Poincaré, Probabilités et Statistiques*, 48(4):909–921, 2012.

[18] Jean Bertoin and Servet Martinez. Fragmentation energy. *Advanced Applied Probability*, 37:553–570, 2005.

[19] Jean Bertoin and Grégory Miermont. The cut-tree of large Galton-Watson trees and the Brownian CRT. *The Annals of Applied Probability*, 23:1469–1493, 2013.

[20] N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular variation.* Cambridge University Press, 1987.

[21] Nicolas Broutin and Minmin Wang. Reversing the cut-tree of the brownian continuum random tree. 2014.

[22] Philippe Chassaing and Bergfinnur Durhuus. Local limit of labeled trees and expected volume growth in a random quadrangulation. *The Annals of Probability*, 34(3):879–917, 2006.

[23] Philippe Chassaing and Gilles Schaeffer. Random planar lattices and integrated superbrownian excursion. *Probability Theory and Related Fields*, (128):161–212, 2004.

[24] Robert Cori and Bernard Vauquelin. Planar maps are well labeled trees. *Canadian Journal of Mathematics*, 33(5):1023–1042, 1981.

[25] Nicolas Curien and Jean-François Le Gall. The Brownian plane. *Journal of Theoretical Probability*, 27(4):1249–1291, 2014.

[26] Nicolas Curien and Grégory Miermont. Uniform infinite planar quadrangulations with a boundary. *Random Structures and Algorithms*. to appear.

[27] Nicolas Curien, Laurent Ménard, and Grégory Miermont. A view from infinity of the infinite planar quadrangulation. 2012.

[28] Daphné Dieuleveut. The cut-tree of Galton–Watson trees converging to a stable tree. *The Annals of Applied Probability*, 25(4):2215–2262, 2015.

[29] R.A. Doney. One-sided local large deviation and renewal theorems in the case of infinite mean. *Probability theory and related fields*, 107(4):451–465, 1997.

[30] Barbara Drossel and Franz Schwabl. Self-organized critical forest fire model. *Phys. Review Letters*, 69:1629–1632, 1992.

[31] Thomas Duquesne. A limit theorem for the contour process of conditionned Galton-Watson trees. *Annals of Probability*, 31(2):996–1027, 2003.

[32] Thomas Duquesne and Jean François Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, 281, 2002.

[33] Thomas Duquesne and Jean François Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probability Theory and Related Fields*, 131:553–603, 2005.

[34] Richard Durrett and Thomas M. Liggett. Fixed points of the smoothing transformation. *Zeitschrift für Wahrscheinlichkeitstheorie*, 64:275–301, 1983.

[35] Christina Goldschmidt and Bénédicte Haas. A line-breaking construction of the stable trees.

[36] Andreas Greven, Peter Pfaffelhuber, and Anita Winter. Convergence in distribution of random metric measure spaces (λ-coalescent measure trees). *Probability Theory and Related Fields*, 145:285–322, 2009.

[37] Misha Gromov. *Metric structures for Riemannian and non-Riemannian spaces*. Birkhäuser, 1999.

[38] Bénédicte Haas, Jim Pitman, and Matthias Winkel. Spinal partitions and invariance under re-rooting of continuum random trees. *Annals of Probability*, 36(5):1790–1837, 2008.

[39] I. A. Ibragimov and Yu. V. Linnik. *Independant and stationnary sequences of random variables*. 1971.

[40] Svante Janson. Random cutting and records in deterministic and random trees. *Random Structures & Algorithms*, 29(2), September 2006.

[41] Harry Kesten. Subdiffusive behavior of random walk on a random cluster. *Annales de l'I.H.P., section B*, 22(4):425–487, 1986.

[42] Maxim Krikun. Local structure of random quadrangulations. 2006. http://arxiv.org/abs/math/0512304.

[43] John Lamperti. A new class of probability limit theorems. *Bulletin of the American Mathematical Society*, 67(3):267–269, 1961.

[44] Jean François Le Gall. Uniqueness and universality of the Brownian map. *Annals of Probability*, 43(4):2880–2960, 2013.

[45] Jean François Le Gall and Yves Le Jan. Branching processes in Lévy processes: The exploration process. *Annals of Probability*, 26:213–252, 1998.

[46] Russell Lyons, Robin Pemantle, and Yuval Peres. Conceptual proofs of l logl criteria for mean behavior of branching processes. *Annals of Probability*, 23(3):1125–1138, 1995.

[47] Jean-François Marckert and Abdelkader Mokkadem. Limit of normalized quadrangulations: the Brownian map. *Annals of Probability*, 34(6):2144–2202, 2006.

[48] A. Meir and J.W. Moon. Cutting down random trees. *Journal of the Australian Mathematical Society*, (11):313–324, 1970.

[49] Grégory Miermont. Self-similar fragmentations derived from the stable tree I: splitting at heights. *Probability Theory and Related Fields*, 127(3):423–454, 2003.

[50] Grégory Miermont. Self-similar fragmentations derived from the stable tree II: splitting at nodes. *Probability Theory and Related Fields*, 131(3):341–375, 2005.

[51] Grégory Miermont. The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Mathematica*, 210(2):319–401, 2013.

[52] Laurent Ménard. The two uniform infinite quadrangulations of the plane have the same law. *Annales de l'I.H.P - Probabilités et Statistiques*, 46(1):190–208, 2010.

[53] Olle Nerman. On the convergence of supercritical general (C-M-J) branching processes. *Zeitschrift für Wahrscheinlichkeitstheorie*, 57:365–395, 1981.

[54] Alois Panholzer. Cutting down very simple trees. *Quaestiones Mathematicae*, 29:211–227, 2006.

[55] Jim Pitman. *Combinatorial stochastic processes*. Springer-Verlag, 2006.

[56] Gilles Schaeffer. Conjugaison d'arbres et cartes planaires aléatoires. 1998. P.h.D. thesis.

[57] William Thomas Tutte. A census of planar maps. *Canadian Journal of Mathematics*, 15:249–271, 1963.

universite PARIS-SACLAY | École doctorale de mathématiques Hadamard (EDMH)

Titre : Coupe et reconstruction d'arbres et de cartes aléatoires

Mots-clefs : arbre de Galton-Watson, arbre brownien, arbre stable, fragmentation auto-similaire, arbre des coupes, quadrangulation uniforme infinie du plan, limite d'échelle, limite locale.

Résumé : Cette thèse se divise en deux parties.
Nous nous intéressons dans un premier temps à des fragmentations d'arbres aléatoires, et aux arbres des coupes associés. Dans le cadre discret, les modèles étudiés sont des arbres de Galton-Watson, fragmentés en enlevant successivement des arêtes choisies au hasard. Nous étudions également leurs analogues continus, l'arbre brownien et les arbres stables, que l'on fragmente en supprimant des points donnés par des processus ponctuels de Poisson. L'arbre des coupes associé à l'un de ces processus, discret ou continu, décrit la généalogie des composantes connexes créées au fur et à mesure de la dislocation. Pour une fragmentation qui se concentre autour de nœuds de grand degré, nous montrons que l'arbre des coupes continu est la limite d'échelle des arbres des coupes discrets correspondants. Dans les cas brownien et stable, nous montrons également que l'on peut reconstruire l'arbre initial à partir de son arbre des coupes et d'un étiquetage bien choisi de ses points de branchement.
Nous étudions ensuite un problème portant sur les cartes aléatoires, et plus précisément sur la quadrangulation uniforme infinie du plan (UIPQ). De récents résultats montrent que dans l'UIPQ, toutes les géodésiques infinies issues de la racine sont essentiellement similaires. Nous déterminons la quadrangulation limite obtenue en ré-enracinant l'UIPQ "à l'infini" sur l'une de ces géodésiques. Cette étude se fait en découpant l'UIPQ le long de cette géodésique. Nous étudions les deux parties ainsi créées via une correspondance avec des arbres discrets, puis nous obtenons la limite souhaitée par recollement.

Title : Cutting and rebuilding random trees and maps

Keywords: Galton–Watson tree, Brownian tree, stable tree, self-similar fragmentation, cut-tree, uniform infinite quadrangulation of the plane, scaling limit, local limit.

Abstract : This PhD thesis is divided into two parts.
First, we study some fragmentations of random trees and the associated cut-trees. The discrete models we are interested in are Galton–Watson trees, which are cut down by recursively removing random edges. We also consider their continuous counterparts, the Brownian and stable trees, which are fragmented by deleting the atoms of Poisson point processes. For these discrete and continuous models, the associated cut-tree describes the genealogy of the connected components which appear during the cutting procedure. We show that for a "vertex-fragmentation", in which the nodes having a large degree are more susceptible to be deleted, the continuous cut-tree is the scaling limit of the corresponding discrete cut-trees. In the Brownian and stable cases, we also give a transformation which rebuilds the initial tree from its cut-tree and a well chosen labeling of its branchpoints.
The second part relates to random maps, and more precisely the uniform infinite quadrangulation of the plane (UIPQ). Recent results show that in the UIPQ, all infinite geodesic rays originating from the root are essentially similar. We identify the limit quadrangulation obtained by rerooting the UIPQ at a point "at infinity" on one of these geodesics. To do this, we split the UIPQ along this geodesic ray. Using a correspondence with discrete trees, we study the two sides, and obtain the desired limit by gluing them back together.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France